MPEG-2 4:2:2

– interoperability and picture-quality tests in the laboratory

Brian Flowers
ex EBU Technical Department

Verification of the correct interoperability of MPEG-2/P@ML encoders and decoders (IRDs) – as supplied by various manufacturers – is an important task which is best carried out by independent third parties, such as the EBU.

This article gives the results of back-to-back interoperability tests carried out in the laboratory by the EBU in June 2001, as a follow-up to satellite tests made in May 2000, when several problems were encountered.

The new results are very encouraging, showing that almost perfect interoperability has now been achieved. Moreover, the picture-quality performance of the latest encoders and decoders shows a noticeable improvement compared with first-generation encoders and decoders.

In May 2000, the EBU carried out MPEG-2/P@ML interoperability tests via satellite, testing encoders and IRDs supplied by nine manufacturers [1]. 11% of the encoder-modulator-IRD combinations produced minor video and/or audio problems and 7% produced unusable video and/or audio. A further 8% of the combinations were not measured for various reasons.

Therefore, it was decided to carry out further tests one year later. However, on this occasion, only the five manufacturers who provide the BISS scrambling system, (Barconet, Nextream, Scientific Atlanta, Scopus and Tandberg), were included in the tests, since the next generation of encoders and IRDs to be implemented on the Eurovision Network will utilize the BISS system.

The tests were carried out back-to-back in the laboratory at EBU headquarters in Geneva. Compared with the earlier tests via satellite, these laboratory tests simplified the procedure and permitted a more detailed analysis of any encoder/decoder problems encountered.

Test procedure

The basic arrangement used for the tests is shown in *Fig. 1*.

(N.B. Barconet provided only a decoder, not an IRD, so this was fed directly with the ASI output of each encoder).

A few minutes of BBC News material from a D5 test tape was first played to check if the video and audio quality, plus the lip-sync, were correct for each encoder-IRD/decoder combination. This was done for each of the three Eurovision Network bit-rates, namely HBR (High Bit-rate), LBR (Low Bit-rate), and ISOG/DSNG. Details of these bit-rates are given in the next section.

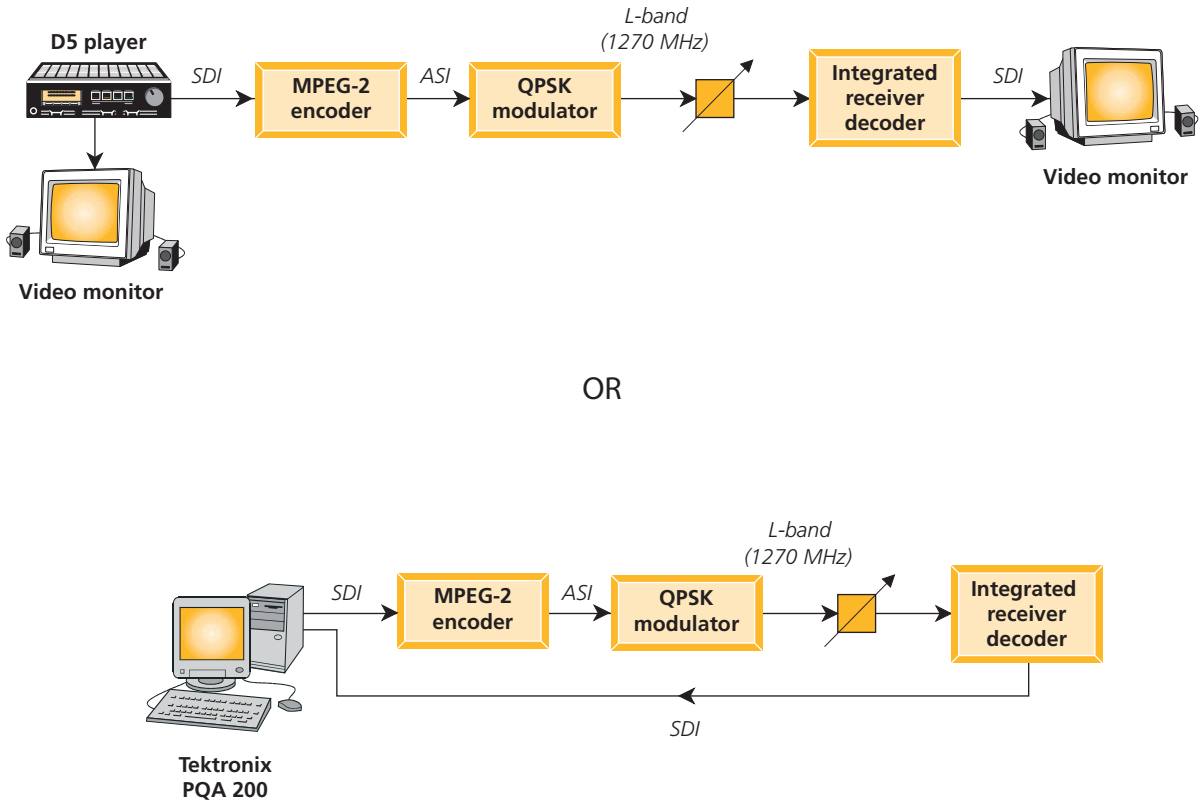

Figure 1
Basic test procedure showing two variants.

The EBU/ITU test sequence "Mobile & Calendar plus header" was then played from the Tektronix PQA 200 and the encoded-decoded signal was analyzed, giving a PQR/YC (Picture Quality Rating, based on luminance and chrominance performance) figure for each encoder-modulator-IRD or encoder-decoder combination, at each bit-rate.

Other aspects of the tests were as follows:

- only 625-line / 25 frames/sec video was tested;
- the encoders were set to IBBP/GoP12, unless shown otherwise in the results tables (at the end of this article), and 4:2:2 coding was selected in all cases;
- the picture resolution was fixed at 720 pixels/line x 576 lines;
- Eurovision Network video PID and audio PIDs (i.e. video: 0x0200, audio 1: 0x1010, audio 2: 0x1020) were selected for each encoder.

Eurovision Network bit-rates:

- **HBR**

Useful bit-rate:	21.503 Mbit/s
Video bit-rate:	19.5 Mbit/s
Audio bit-rate:	2 x 384 kbit/s
Bit-rate with RS:	23.333 Mbit/s
Symbol rate of QPSK modulation with 7/8 FEC:	13.333 Msymb/s

- **LBR**

Useful bit-rate:	10.7515 Mbit/s
Video bit-rate:	9.0 Mbit/s
Audio bit-rate:	2 x 384 kbit/s
Bit-rate with RS:	11.6665 Mbit/s
Symbol rate of QPSK modulation with 7/8 FEC:	6.6666 Msymb/s

- **ISOG/DSNG**

Useful bit-rate:	8.448 Mbit/s
Video bit-rate:	7.0 Mbit/s
Audio bit-rate:	2 x 256 kbit/s
Bit-rate with RS:	9.167 Mbit/s
Symbol rate of QPSK modulation with 3/4 FEC:	6.1113 Msymb/s

The Tektronix PQA 200

This equipment was developed by Tektronix and Sarnoff in the USA to provide objective measurement of digital video quality.

The EBU's Tektronix PQA 200 equipment has software version 3.0.

Several perfect 270 Mbit/s test sequences are stored on a hard disk in the generator. These sequences include header stripes, which identify each frame, thereby permitting a frame-by-frame comparison by the analyzer of the original signal and the stored encoded-decoded signal.

The encoded-decoded signal is automatically corrected in terms of luminance and chrominance amplitudes before being analyzed, thereby avoiding misleading results due to level disparities.

The analyzer takes several minutes to compare the original signal with the stored encoded-decoded signal, using special filters to ensure that the measured degradation correlates well with subjective assessment of the degradation carried out in accordance with ITU-R BT.500 [2].

The PQR values are classified as follows:

- A PQR of one unit corresponds to a *just perceptible* increase in degradation, according to the ITU-R BT.500 subjective test procedure. (However, a degradation corresponding to a PQR value <3 is not perceptible to most people.)

Abbreviations

ASI	Asynchronous Serial Interface
BISS	(EBU) Basic Interoperable Scrambling System
DSNG	Digital Satellite News Gathering
FEC	Forward Error Correction
IRD	Integrated Receiver/Decoder
HBR	High Bit-Rate
GoP	Group of Pictures
ISOG	Inter-Union Satellite Operations Group
ITU	International Telecommunication Union
ITU-R	ITU - Radiocommunication Sector
LBR	Low Bit-Rate
MPEG	Moving Picture Experts Group
MPEG-2/P@ML	(MPEG-2) Professional Profile at Main Level
PID	(MPEG) Packet Identification Number
PQR	Picture Quality Rating
QPSK	Quadrature (Quaternary) Phase-Shift Keying
RS	Reed Solomon
SDI	Serial Digital Interface
YC	The luminance (Y) and chrominance (C) signals of a colour television system

- PQR values in the range 3 - 6 correspond to *perceptible but not annoying* degradation.
- PQR values in the range 6 - 9 correspond to *slightly annoying* degradation.
- PQR values in the range 9 - 12 correspond to *annoying* degradation.
- PQR values >12 correspond to *very annoying* degradation

PQR values of 9 and above are not acceptable for contribution links.

(Note that this classification is not an officially approved scale but only an attempt to correlate PQR values with the long established five point ITU quality scale, based on our experience with digital contribution links).

Figure 2
Off-screen image of the "Mobile and Calendar" test sequence.

The "Mobile & Calendar" test sequence has medium to high criticality, so it provides quite a good test of encoder performance. However, it does not test all aspects of encoder performance, since it does not contain a scene-change cut, for example.

5. Results

- No interoperability problems were encountered at LBR and ISOG/DSNG video bit-rates.

- At HBR, one minor interoperability problem was encountered.

  The Nextream encoder (DBE 4110) feeding the Tandberg/Alteia decoder gave a higher than expected PQR of 4.59, which is about 2 PQR units higher than the PQRs for other combinations of encoder-decoder.

  Having encountered this problem in the May 2000 tests, we knew that it was due to a buffer mismatch, so a test was made with the Nextream encoder set to "low delay" encoding, namely IP/GoP12, which gave a good PQR result of 2.72.

  The Tandberg/Alteia decoders are normally selected to "short buffer" for use on the Eurovision Network, because the first-generation NDS encoders and decoders had only the short buffer available.

  This situation can lead to the above buffer mismatch problems.

  When all first-generation NDS equipment is finally decommissioned, the second-generation Tandberg encoders and decoders can be selected to "normal buffer", thereby eliminating this buffer mismatch problem.

  In view of this problem, the Tandberg/Alteia decoder was tested with both "short buffer" and "normal buffer", giving no interoperability problems in the latter case.

  The Barconet Polaris encoder gave PQR values slightly higher than the other encoders. However, a new encoder module, which unfortunately was not available in time for these tests, is expected to give better results.

- Complete results for the three bit-rates are given in *Tables 1, 2* and *3* (at the end of this article).

  It can be seen that the valid PQR/YC values at HBR are in the range 2.44 to 2.95, so degradation at this bit-rate is imperceptible for all combinations of encoder-decoder.

  The first generation of MPEG-2/P@ML encoders (NDS 3000) gave a PQR/YC of 4.58 at HBR, so an improvement of about 2 PQR units has been achieved with the second generation of encoders.

  The PQR/YC values at LBR are in the range 3.87 to 4.46, and in the range 4.35 to 5.03 for ISOG/DSNG bit-rates, all these values corresponding to *perceptible but not annoying* degradation.

Brian Flowers studied Engineering at the University of Southampton, UK. In 1960, after serving two years in the Royal Air Force (RAF), he joined BBC Television News. Then, in 1962, he was detached to the Eurovision Control Centre (EVC) in Brussels.

Mr Flowers worked at all levels of responsibility in the EVC – technician, supervisor, engineer-in-charge, and as the project leader for the new EVC in Geneva.

Prior to his retirement in 1997, Brian Flowers was a Senior Engineer in the Transmission Technology division of the EBU Technical Department, Geneva.

The corresponding PQR/YC for first-generation NDS encoders were 5.47 at LBR and 5.93 for ISOG/DSNG respectively, so an improvement of about 1.5 PQR units at LBR and 1 PQR unit at ISOG/DSNG bit-rate has been achieved in the second-generation equipment.

Conclusions

- No interoperability problems were encountered at the Low Bit-rate and the ISOG/DSNG bit-rate.
- One minor interoperability problem was encountered at the High Bit-rate. This was due to a well-known buffer-size incompatibility problem.
- The encoders largely determine the PQR values for a given bit-rate, whereas the decoders have very little influence on the PQR values.
- There was no significant variation in performance between the various encoders and decoders.
- Picture degradation for a single encoding-decoding operation at HBR is imperceptible, whereas at LBR and the ISOG/DSNG bit-rate, the degradation is *perceptible but not annoying*.
- The performance of the second-generation encoders is better than that of first-generation encoders.

  The improvement – based on the Tandberg 3000 and Tandberg 5000 encoders, for example – is about 2 PQR units at HBR, 1.5 PQR units at LBR, and 1 PQR unit at the ISOG/DSNG bit-rate.

Bibliography

[1] B. Flowers: **MPEG-2 4:2:2 — interoperability tests via satellite**
EBU Technical Review No. 285, December 2000.

[2] ITU-R BT.500: **Methodology for the subjective assessment of the quality of television pictures**
http://www.itu.int/rec/recommendation.asp?type=folders&lang=e&parent=R-REC-bt.500

Tables of results

Table 1
EBU/HBR — IBBP/GoP12 *(except when shown otherwise)*
(Mobile and Calendar plus header)
Tektronix PQA 200. PQR/YC values are shown below.

Encoder \ IRD	Tandberg Alteia		Scopus (IRD 2800)	Scientific Altanta (D 9224)	Nextream (DBD 4432)	Barconet (Stellar)
	Short buffer	Normal buffer				
Tandberg 5000 (E 5611)	2.57	2.60	2.59	2.58	2.58	2.57
Scopus (E 1000)	2.46	2.50	2.50	2.46	2.48	2.46
Scientific Atlanta (D 9150)	2.46	2.45	2.46	2.46	2.44	2.46
Nextream (DBE 4110) (IP/GoP12) Low delay	4.59 [a] 2.72	2.95 2.75	2.85 2.77	2.71 n.m. [b]	2.75 2.73	2.80 2.73
Barconet (Polaris)	2.75	2.74	2.77	2.75	2.74	2.73

a. Poor PQR figure due to encoder-decoder buffer-size mismatch.
b. Not measured.

PQR range is: 2.44 – 2.95

Table 2
EBU/LBR — IBBP/GoP12
(Mobile and Calendar plus header)
Tektronix PQA 200. PQR/YC values are shown below.

Encoder \ IRD	Tandberg Alteia	Scopus (IRD 2800)	Scientific Altanta (D 9224)	Nextream (DBD 4432)	Barconet (Stellar)
	Short buffer				
Tandberg 5000 (E 5611)	3.90	3.92	3.94	3.94	3.97
Scopus (E 1000)	3.93	3.94	4.12	4.12	4.13
Scientific Atlanta (D 9150)	4.13	4.14	4.12	4.13	4.14
Nextream (DBE 4110)	3.89	3.90	3.87	3.90	3.90
Barconet (Polaris)	4.46	4.43	4.43	4.43	4.36

PQR range is: 3.87 – 4.46

Table 3
ISOG/DSNG — IBBP/GoP12
(Mobile and Calendar plus header)
Tektronix PQA 200. PQR/YC values are shown below.

IRD / Encoder	Tandberg Alteia Short buffer	Scopus (IRD 2800)	Scientific Altanta (D 9224)	Nextream (DBD 4432)	Barconet (Stellar)
Tandberg 5000 (E 5611)	4.51	4.56	4.35	4.59	4.50
Scopus (E 1000)	4.58	4.60	4.65	4.90	4.88
Scientific Atlanta (D 9150)	4.68	4.67	4.66	4.89	4.89
Nextream (DBE 4110)	4.54	4.54	4.55	4.51	4.52
Barconet (Polaris)	4.95	4.94	4.99	5.03	4.92

PQR range is: 4.35 – 5.03