MPEG-2 ENCODING AND MOST COMMON PROFILES: 4:2:0 (MP@ML) AND 4:2:2

The advantages of digital TV transmission, in comparison with analogue, considering both microwave links and actual broadcasting, are notable and evident from the following:

- more TV programmes may be transmitted in a given RF spectrum (typically at least four times as many)
- lower transmission power will cover the same distance (or same power will give greater immunity to noise and interference)
- better transmission quality
- the potential for using Single-Frequency terrestrial broadcasting Networks - SFN (using COFDM modulation)
- the possibility of mobile reception (using COFDM modulation)
- the potential for simultaneous transmission of auxiliary data

Digital TV transmission needs digital audio and video signals. These signals may be originated digitally (with all-digital cameras and studio mixers) or, more usually, they may be produced by digitally encoding available analogue signals.

Uncompressed digital video and audio signals have high data rate - typically one programme requires a bit-rate of 270Mbit/s. Normally, the Serial Digital Interface (SDI) is used for this type of signal (with 75 ohm BNC coaxial connectors).

If uncompressed data were to be transmitted as is, the occupied RF bandwidth would be much greater than in the analogue case. It is necessary, therefore, to compress such data to a lower rate, making it suitable for transmission over microwave links and for distribution or broadcasting to viewers.

This compression is required ideally, not to degrade the quality of the video or audio signals. The designated international coding standard for this purpose is MPEG-2 (Motion Picture Expert Group version 2) which is able to compress a TV programme from 270Mbit/s to only 5 or 6Mbit/sec while maintaining excellent quality characteristics. Compression to less than 4Mbit/s is possible but quality will be compromised.

The following compression techniques are used to encode TV pictures:

- Human visual perception is more sensitive to luminance than chrominance. Less information (data) about the colour is therefore transmitted.
- Adjacent areas within the picture often have pixels with the same luminance and chrominance values. During encoding these are combined so as to transmit less data.
- Only the differences between one picture frame and the next are transmitted. This process is carried out several times over a Group Of Pictures (GOP) before eventually transmitting a complete frame again.

So GOPs – Groups of Pictures – are made up from three different kinds of information frames:

- I-frame: the complete image or picture frame (the largest in terms of the data transmitted)
- P-frame: the differences between an actual and the previous I or P-frame (smaller than an I-frame)
- B-frame: the differences between the previous and the following I or P frames (the smallest frame, but which cannot be repeated too many times).

Usually GOPs are constituted with one I-frame, some P-frames and, possibly, some B-frames. They should not be too long because should an error occur, it would be perpetuated. Furthermore, a decoder requires a complete picture (I-frame) to begin decoding, so has to wait for the start of a GOP.

One of the most usual and efficient GOP structures is 12 frames long and is constituted as follows: IBBPBBPBBPBB.

The most common encoding data profiles are 4:2:0 (Main Profile @ Main Level or MP@ML) and 4:2:2. We list below the properties, advantages and uses of each:

- **4:2:0** – The video is encoded with a ratio of 4 data elements for luminance to 2 for chrominance.

  ADVANTAGES:

  - This encoding ratio matches the visual perception characteristic
  - Optimum performance, particularly for low Bit-Rate transmission

  USES:

  - Broadcasting (the profile used in both terrestrial and satellite broadcasting)
  - Contribution and Distribution networks
  - Intra-studio links between analogue and digital mixers

- **4:2:2** – The Video is encoded with a ratio of 4 data elements for luminance to 4 for chrominance

  ADVANTAGES:

  - Slightly better performance than 4:2:0 profile, but only when the Bit-Rate is over 10MBit/s

  USES:

  - Intra-studio links between digital mixers

The above summary is supported in the respected EBU Technical Review, its autumn-1999 issue, reporting on a series of tests carried out by the Swedish Television Authority (SVT). Comparative tests, according to ITU recommendations, were based on subjective evaluation by a group of observers, viewing digital TV pictures after encoding at 2, 3, 4 or 5Mbit/s. The results established that, for the each Bit-Rate, 4:2:0 was preferable to 4:2:2 encoding.

Also, Mr Al Kovalick (Technical Director of Pinnacle Systems Inc) says in an article published by BroadcastPapers (© 2001-2002) that a video sequence encoded with 4:2:0 at 10Mbit/s has the same quality when encoded using 4:2:2 profile, but at 13Mbit/s!

During our tests (at ABE Elettronica), we found some advantages using 4:2:2 encoding, but only at a Bit-Rate over 10Mbit/s. It is preferable to use the 4:2:0 profile for Bit-Rates under 10Mbit/s, especially for picture sequences with high motion content.

Please note that it is extremely difficult to detect quality differences on picture sequences encoded at over 10Mbit/s, since the quality is already so high that differences are very difficult to perceive.

The limited advantages of the 4:2:2 profile compared with 4:2:0 with Bit-Rates over 10Mbit/s disappear if the source signals are analogue in origin and converted to digital.

So, considering that it is now unusual (and expensive) to use Bit-Rates of 15-20Mbit/s just for a single programme, the encoding profile used is nearly always 4:2:0 (MP@ML)

We would like to mention here some of the most common settings for an MPEG-2 encoder (4:2:0 profile):

- Video resolution: Full D1; 3/4 D1; 2/3 D1; 1/2 D1; SIF; QSIF (One has to choose the most appropriate setting keeping in mind the available Bit-Rate and required encoding quality according to the content).
- Resolution of displayed pictures: 720 x 576 pixel, max for PAL, and 720 x 480 pixel, max for NTSC. Higher resolution can produce better definition, but at the expense of higher Bit Rate.
- Group of Pictures (GOP) structure: the number and sequence of encoded I, P, B frames.
- Encoding Bit-Rate: up to 15MBit/s.
- Output Transport Stream Bit-Rate: has to be equal to or higher than the total from the video and audio encoding, plus the data-tables. The difference between the real encoding Bit-Rate and the output Transport Stream Bit-Rate is made up by filling with null packets (bit-stuffing).
- Audio sampling frequency (32 or 44.1 or 48 kHz): the higher the sampling frequency, the better the transmission quality, but the higher the necessary Bit-Rate.
- Video, Audio and, possibly, Data PIDs (Program Identifiers): these have to be set avoiding duplication so as not to be in conflict with other PIDs with which they may be multiplexed.
- Filter settings: in the case of encoders with composite video input is possible to choose 'comb' or 'notch' filters to separate chrominance and luminance. Other kinds of filter are useful to reduce noise (for example, in the case of low Bit-Rate to avoid transmitting noise instead of real pictures).

The above settings are only some of those available on a typical encoder. In practice, for example, on the MPEG-2 (4:2:0) encoder produced by ABE Elettronica, it is possible to modify some 200 parameters, although naturally many of these are inter-dependent. To help the user, the ABE encoder has four different factory set-up configurations and four others which are user-configurable.

Finally, some considerations about the testing of MPEG-2 encoders:

- Encoding Quality, as such, is not easily measurable. The usual method of assessment is to make comparisons of picture sequences with subjective evaluation and/or expert viewing, rather than using the few test and measuring sets available which may not match human quality perception.
- Supporting chosen method of quality evaluation (but not replacing it), it is possible also to measure linear and non-linear video distortions and noise. It should be kept in mind, however, that many of the results will depend on the encoder settings as well as on its 'quality' and detailed algorithms. Moreover, the effective luminance pass-band could be less than 3 MHz, depending on the filters used.

ADVANCED BROADCASTING ELECTRONICS

ABE ELETTRONICA S.p.A.
Via Leonardo da Vinci, 92
24043 ● CARAVAGGIO (BG) ● Italy

Tel. +39-0363-351.007 ● FAX +39-0363-50.756
www.abe.it