MPEG2

4C8: Digital Media Processing

Ussher Assistant Professor François Pitié

2017/2018

Department of Electronic & Electrical Engineering , Trinity College Dublin

adapted from original material written by Prof. Anil Kokaram.

Overview

In any multimedia transmission system that involves compression the following issues become important.

Overview

Data Type. Multimedia data sources can be pictures, audio and text. Different compression techniques are needed for each data type. Each piece of data has to be identified with unique codewords for transmission.

Overview

Sequencing. The compressed data from each source is scanned into a sequence of bits. This sequence is then packetised for transport. The problem here is to identify each different part of the bitstream uniquely to the decoder, *e.g.* header information, DCT coefficient information.

Overview

Multiplexing The audio and video data (for instance) has to be decoded at the same time to create a coherent signal at the receiver. This implies that the transmitted elementary data streams should be combined so that they arrive at the correct time at the decoder. The challenge is therefore to allow for identifying the different parts of the multiplexed stream and to insert information about the timing of each elementary data stream.

Overview

Media. The compressed and multiplexed data has to be stored on some digital storage and then later (or live) broadcast to receivers across air or other links. Access to different Media channels is governed by different constraints and this must somehow be allowed for in the standards description.

Overview

Errors. Errors in the received bitstream invariably occur. The receiver must cope with errors such that the system performance is robust to errors or it degrades in some graceful way.

bandwidth

Bandwidth. The bandwidth available for the multimedia transmission is limited. The transmission system must ensure that the bandwidth of the bitstream does not exceed these limits.

The problem of tuning the encoder to meet these expectations is called **Rate Control** and applies both to the control of the bitrate of the elementary data streams and the multiplexed stream.

Overview

Multiplatform. The coded bitstream may need to be decoded on many different types of device with varying processor speeds and storage resources. It would be interesting if the transmission system could provide a bitstream which could be decoded to varying extents by different devices. Thus a low capacity device could receive a lower quality picture than a high capacity device that would receive further features and higher picture quality. This concept applied to the construction of a suitable bitstream format is called Scalability.

Profiles and Levels

The MPEG standards have been designed to cope with a wide variety of picture formats and frame rates. However not all decoders and encoders will be able to cope with all possible combinations of input picture formats. Therefore a hierarchy of data formats was specified such that the capability of a Codec could be defined in terms of a combination of various input data options allowed by the standard.

The standard does not define encoders or decoders themselves, it only defines the format of a bitstream. This format implicitly defines much of the decoder but creation of the bitstream (the encoder) is outside the standard.

Profiles

Profile	Supported Features
HIGH	Supports the SPATIAL Scalable Profile + 3 layers with SNR and Spatial Scalable coding modes 4:2:2 YUV picture format
SPATIAL	Supports the SNR Scalable Profile + Spatial Scalable coding modes (2 layers) 4:0:0 YUV
MAIN	Supports SIMPLE profile + B-picture prediction modes
SIMPLE	Supports coding progressive and interlaced video random access I, P-picture prediction modes 4:2:0 YUV

Levels

Level	Upper Bound on Parameters
HIGH	1920 samples/line , 1152 lines/frame 60 frames/sec 80 Mbits/sec
HIGH 1440	1440 samples/line, 1152 lines/frame 60 Frames/sec 60 Mbits/sec
MAIN	720 samples/line, 576 lines/frame 30 Frames/sec 30 Mbits/sec
LOW	352 samples/line, 288 lines/frame 30 frames/sec 4 Mbits/sec

Layers

The MPEG-2 bitstream is organised with a strict hierarchy of layers as follows:

1. The **Block** Layer: contains the DCT coefficients for a single block.
2. The **Macroblock** Layer: contains a 16×16 block of pixels aranged in 4 8×8 blocks. It is the motion compensation unit in that it is at this layer that vectors are associated with blocks.
3. The **Slice** Layer. A row of Macroblocks can be sliced into different groups of macroblocks. Slices are the MPEG-2 solution to mitigate the propagation of errors.
4. The **Picture** Layer. This section contains all the encoded data referring to a single I, P or B frame.
5. The **GOP** Layer. This delineates a unit of frames that can be decoded independantly.
6. The **Sequence** Layer. This contains an entire video sequence with thousands of frames.

Block Layer

As with JPEG, the MPEG committee has selected default values for the quantisation matrices for the 8×8 DCT block.

The Default Intra Quantisation Matrix:

$$\begin{pmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 47 & 49 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{pmatrix}$$

Block Layer

As compression for the Inter-coded images is done on the DFD instead of the actual image, we should employ a different quantisation matrix.

The Default Inter Quantisation Matrix:

$$\begin{pmatrix} 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\ 16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \end{pmatrix}$$

Macro Blocks

In MPEG-2 motion estimation and compensation is carried out on blocks of 16×16 and NOT 8×8. Thus only ONE vector is transmitted per 16×16 macroblock for P Pictures and up to TWO for B Pictures.

Motion vectors are encoded using difference coding. (ie. only the difference between a macro block motion vector and the previsouly estimated motion vector on the macroblock to the left is sent for entropy coding.)

In MPEG-2 motion vectors are quantised to ± 0.5 pixel accuracy.

Multiplexing

Audio and video streams are divided into packets that are interleaved by the encoder.

To solve multiplexing problems, synchronisation between packets is achieved by the use of time stamps.

The **Decoding Time Stamp (DTS)** flag tells the decoder when to decode a packet. The **Presentation Time Stamp (PTS)** flag tells the decoder when to pass the decoded frame to the output device for display.

Going Further

The book *Digital Video: An Introduction to MPEG-2, Barry Haskell, Atul Purui and Arun Netravali, Chapman and Hall 1997* is an excellent overview of MPEG2.

Prof. Inald Lagendijk at the Delft University of Technology in The Netherlands has an excellent MPEG2 demo tool called VCDEMO which he uses to teach an image and video coding course. It is available free on the web.