MPEG's parts (layers)

- part 1 - Describes synchronization and multiplexing of video and audio
- part 2 - Describes compression of video signals
- part 3 - Describes compression of audio signals
- part 4 - Compliance Testing

© Klein S. T. and Wiseman Y.

MPEG

- MPEG is Moving Picture Experts Group
- On 1992 MPEG-1 was the standard, but was replaced only a year after by MPEG-2.
- Nowadays, MPEG-2 is gradually replaced by MPEG-4, which was finalized on October 98.
- MPEG-3 was integrated into MPEG-2.

© Klein S. T. and Wiseman Y.

The video compression

- The images are in color and converted into YUV space.
- The resolution of the Y (luminance) channel is twice as good as the U and V channels. Each block of U and V is 8x8, while each block of Y is 16x16.
- The blocks go through DCT, Quantization and Entropy Encoder, like JPEG.

© Klein S. T. and Wiseman Y.

File Format

- The file format is designed to be independent of any particular delivery protocol
- The MP4 file format is composed of object-oriented structures called 'atoms'. A unique tag and a length identify each atom.
- Part 1 is made up of such atoms.
- The other parts can be located elsewhere. They can be in media data atoms, another file, or even on the web.
- The rates of the video and the audio don't have to be equal.

© Klein S. T. and Wiseman Y.

Sequence of frames

- The sequence of decoded frames usually goes like:

IBBPBBPBBPBBIBBPBBPB...

- The compressed data stream ends up looking like:

0312645...

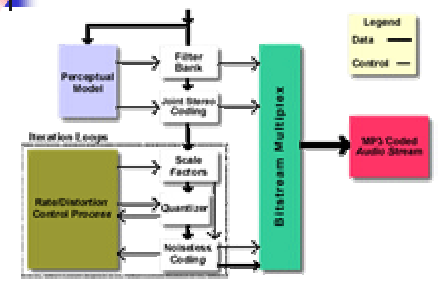
© Klein S. T. and Wiseman Y.

Types of frames

- I - Intra frame - a frame coded as a still image, not using any past history.
- P - Predicted frame - predicted from the most recently reconstructed I or P frame.
- B - Bidirectional frame - predicted from the closest two I or P frames, one in the past and one in the future.

© Klein S. T. and Wiseman Y.

The audio compression



 © Klein S. T. and Wiseman Y.

Frames distribution

- Usually, the size of a P frame is about 70% of an I frame and the size of a B frame is about 35% of an I frame .
- There are no more than 12 frames from I to I. This is based on a random access requirement and error recovery.
- The ratio of P's to B's is based on experience.

 © Klein S. T. and Wiseman Y.

Quantization and Coding

- Scale Factors - Divides by a constant factor.
- Quantization - Larger values are automatically coded with less accuracy and some noise shaping is already built into the quantization process.
- noiseless coding - The quantized values are coded by Huffman coding. As a specific method for entropy coding, Huffman coding is lossless. Thus is called noiseless coding.

 © Klein S. T. and Wiseman Y.

Filtering

- Filter Bank - Consists of a polyphase filter bank and a Modified Discrete Cosine Transform (MDCT).
- Perceptual model - Determining the quality of a given encoder implementation. It uses either a separate filter or a combination.
- Joint stereo coding - Takes advantage of the fact that both channels of a stereo channel pair contain far the same information.

 © Klein S. T. and Wiseman Y.

Rate/Distortion Control

- If the number of bits resulting from the coding operation exceeds the number of bits available to code a given block of data, this can be corrected by adjusting the global gain to result in a larger quantization step size.
- This operation is repeated with different quantization step sizes until the resulting bit demand for Huffman coding is small enough.
- The loop is called rate loop because it modifies the overall coder rate until it is small enough.

 © Klein S. T. and Wiseman Y.

Low scale factors

- When a low scale factor is used, a bit-slicing scheme is applied to the quantized spectral data.
- First the quantized spectral values are grouped into frequency bands. Each of these groups contains the quantized spectral values in their binary representation.
- The bits of a group are processed in slices according to their significance. Thus first all most significant bits (MSB) of the quantized values in a group are processed, etc.
- These bit-slices are then encoded using an arithmetic coding scheme to obtain entropy coding with minimal redundancy.

 © Klein S. T. and Wiseman Y.

Independent audio compression

- Layer 3 can be use independently, used to be called MP3.
- Usually, MP3 gives almost the original sound quality. The changes are unnoticeable.
- The data reduction is typically 1:12 for a good stereo CD.
- Telephone lines have much lower quality. Hence the data reduction is drastically improved - about 1:96.

© Klein S. T. and Wiseman Y.

Bitstream Multiplex

- The average amplitude is taken from the filter bank.
- The average frequency is taken from the Joint Stereo Coding.
- The Huffman encoded data actually represents the changes in amplitude and frequency.
- This signal representation permits speed and pitch change functionality by simple parameter modification in the decoder.

© Klein S. T. and Wiseman Y.