Claims

1. Encoder for encoding secondary media data comprising metadata or control data for primary media data, the encoder comprising:

a grouper configured for grouping a bitstream of secondary media data to form grouped secondary media data, representing data words;

a reference signal generator for generating a reference pattern indicating a reference amplitude or a predetermined timing instant in the primary media data; and

a stream builder comprising a pulse shaping filter to low-pass filter the data words and the reference pattern to obtain data pulses comprising a length of more than one sample of a predetermined sample rate, wherein an amplitude of a data pulse is weighted according to the data words and the reference pattern, and wherein the pulse shaping filter is configured to add up consecutive weighted data pulses at instants of the predetermined sample rate to obtain a stream of digital words,

wherein the encoder is configured to output the stream of digital words as encoded secondary media data.

2. Encoder according to claim 1, wherein the encoder is configured to adding redundancy by a digital modulation.

3. Encoder according to claim 1 or 2, wherein a digital modulation is so that two or more bits of the secondary media data are transmitted per digital word of the stream of digital words.

4. Encoder according to any one of the preceding claims, wherein the encoder is configured to output the stream of digital words so that the stream of digital words is transmittable over a PCM audio channel.

5. Encoder according to any one of the preceding claims, wherein the encoder is configured to output a further stream of digital words, the further stream of digital

words representing the primary media data, the further stream being separate from the stream of digital words.

6. Encoder according to claim 5, wherein the primary media data are audio data, and wherein the secondary media data are metadata for the audio data or control data for the audio data.

7. Encoder according to claim 5 or 6, wherein the encoder is configured to output the stream of digital words and the further stream of digital words so that the further stream of digital words is transmittable over a first audio PCM channel and so that the stream of digital words is transmittable over a second audio PCM channel being different from the first audio PCM channel.

8. Encoder of claim 5 or 6 or 7, wherein each of the digital words of the further stream representing the primary media data has a predefined number of bits being greater than 8 bits and smaller than 32 bits, and wherein each of the digital words of the stream of digital words has the predetermined number of bits as well.

9. Encoder according to any one of the preceding claims, wherein a digital modulation is a pulse amplitude modulation.

10. Encoder according to claim 1, wherein the pulse shaping filter is a raised cosine shaping filter.

11. Encoder according to any one of the preceding claims, wherein the encoder is configured to output a video stream representing a sequence of video images, and wherein the encoder is configured to output the stream of digital words so that the control data or meta data of the secondary media data related to a certain video image are related to the certain video image.

12. Encoder according to claim 11, wherein the encoder is configured to output the stream of digital words as a first stream of digital words associated to a first video image of the sequence of video images, and to output the stream of digital words as a second stream of digital words associated to a second video image of the sequence of video images, wherein the first and second digital words are identical to each other.

13. Encoder according to any one of the preceding claims,

wherein the encoder is configured to generate the digital words, the digital words having 12 to 28 bits, or wherein the digital words are sampled at a sampling rate of between 30 kHz to 55 kHz, or wherein the digital words have a dynamic range of 70 to 160 dB, or have a nominal signal level of -20 dB RMS full scale.

14. Encoder according to any one of the preceding claims, wherein the encoder is configured to use an upper frequency for band limiting the secondary media data being between 15 kHz to 27,5 kHz for a sampling rate between 30 kHz to 55 kHz.

15. Encoder according to any one of the preceding claims, the encoder comprising:

a mapper configured for mapping a group of secondary media data comprising a first number of bits into a data word comprising a second number of bits being greater than the first number of bits and wherein the grouped secondary media data is aligned with a gap to a most significant bit or a least significant bit of the data word.

16. Encoder according to claim 1,

wherein the data word comprises the grouped secondary media data, being sign-extended to a most significant bit and a gap between the grouped secondary media data and a least significant bit padded with zeroes, and wherein the reference pattern comprises the same structure as the data word.

17. Encoder according to any one of the preceding claims,

wherein the pulse shaping filter is configured to obtain zero points at instants of the predetermined sample rate of the data pulses;

wherein the stream builder is configured to build the stream representing the encoded secondary media data using the reference pattern and the plurality of data words such that zero points of a data pulse are aligned with a maximum of a

further data pulse to obtain an inter-symbol-interference-free stream representing the encoded secondary media data.

18. Encoder according to any one of the preceding claims, wherein the pulse shaping filter comprises a cutoff frequency of less than 1.5 times of a sampling frequency of the primary media data.

19. Encoder according to any one of the preceding claims, wherein the reference signal generator is configured to generate a grouped reference pattern comprising a first number of bits and wherein the reference signal generator is further configured to map the grouped reference pattern into a data word comprising a second number of bits being greater than the first number of bits; or

wherein a mapper is configured to map a grouped reference pattern comprising a first number of bits into a data word comprising a second number of bits being greater than the first number of bits.

20. Encoder according to any one of the preceding claims, wherein the encoder is configured to output the encoded secondary media data as the stream of digital words as a control track and to output up to 15 channels of the primary media data as audio tracks, wherein the control track and the audio tracks are formed in accordance with the AES 3 standard.

21. Decoder for decoding a media signal comprising a received stream of digital words representing encoded secondary media data comprising metadata or control data for primary media data and a reference pattern indicating a reference amplitude or a predetermined timing instant in the primary media data, the decoder comprising:

a reference pattern analyzer for analyzing the reference pattern of the encoded secondary media data, wherein the reference pattern analyzer is configured to determine the reference amplitude of the reference pattern or to determine a predetermined timing instant in the primary media data;

a signal manipulator for manipulating the encoded secondary media data in accordance with the analyzed reference pattern and a computed reference pattern to obtain secondary media data, wherein the signal manipulator is configured, in order to obtain a bitstream from the recovered secondary media data, to

calculate and compensate an amplification or attenuation or a DC offset applied to the stream of digital words; or

to convert a sample rate associated with the digital words, according to the predetermined timing instant of the primary media data indicated in the reference pattern, to a predetermined sample rate

.

22. Decoder according to claim 21, wherein the recovering comprises manipulating the received stream of digital words with respect to amplitudes represented by the received digital words.

23. Decoder according to claim 21 or 22, wherein the media signal additionally comprises a further received stream of digital words representing encoded primary media data, the further received stream being separate from the received stream, and

wherein the decoder is configured to process the primary media data represented by the further received stream using the metadata or control data represented by the bitstream.

24. Decoder according to any one of claims 21 to 23 comprising:

wherein the reference pattern analyzer is configured to determine the amplitude of the reference pattern by averaging central four samples of a PCM audio signal during a pulse period and during a pulse period in the reference pattern; or

wherein the decoder further comprises a signal processor for processing the primary media data according to the encoded secondary media data to obtain a decoded media signal.

25. Decoder according to any one of claims 21 to 24, wherein the reference pattern analyzer comprises:

a timing instant determiner configured to determine the predefined timing instant of the primary media data in the reference pattern in terms of samples of a sample rate;

an upsampler configured to upsample a range around the determined timing instant to determine an exact position of a predetermined timing instant;

a sampling accumulator configured to determine an exact position of the digital words within the stream of digital words to obtain an actual sample rate associated to the digital words being different from a predetermined sample rate.

26. Decoder according to any one of claims 21 to 25, wherein the reference pattern analyzer comprises a gain factor calculator to calculate an amplification or attenuation factor according to the amplitude of the reference pattern and the amplitude of the computed reference pattern and wherein the signal manipulator comprises a multiplier configured to amplify or attenuate the data words according to the amplification or attenuation factor to obtain gain compensated data words.

27. Decoder according to any one of claims 21 to 26,

wherein the reference pattern analyzer comprises an amplitude detector to determine an amplitude of a reference pattern and a further amplitude of the reference pattern;

wherein the reference pattern analyzer further comprises an offset compensation unit configured to calculate an offset of the encoded secondary media data according to a drift of the amplitude of the reference pattern and the further amplitude of the reference pattern, wherein the signal manipulator comprises an adder configured to add the offset of the encoded secondary media data to the encoded secondary media data to obtain offset compensated encoded secondary media data.

28. Decoder according to any one of claims 21 to 27,

wherein the signal manipulator comprises a demapper configured to demap secondary media data comprising a first number of bits from data words comprising a second number of bits being greater than the first number of bits; or

wherein the signal manipulator comprises an ungrouper configured to ungroup a group of secondary media data comprising a first number of bits to obtain the bitstream of decoded secondary media data.

29. Media signal comprising:

a stream of digital words representing encoded secondary media data comprising metadata or control data for primary media data and a reference pattern indicating a reference amplitude or a predetermined timing instant in the primary media data;

the stream of digital words comprising data pulses comprising low-pass filtered data words and the reference pattern low-pass filtered by a pulse shaping filter, the data pulses comprising a length of more than one sample of a predetermined sample rate, wherein an amplitude of a data pulse is weighted according to the data word or the reference pattern, and wherein the pulse shaping filter is configured to add up consecutive weighted data pulses at instants of the predetermined sample rate to obtain the stream of digital words.

30. Media signal according to claim 29, wherein the digital words comprise pulse shaped and low-pass filtered secondary media data.

31. Media signal according to claim 29 or 30, wherein the secondary media data comprises the reference pattern and a plurality of data words.

32. Media signal according to one of claims 29 to 31, wherein the secondary media data is mapped into data words with a gap to the most significant bit of the data word or to the least significant bit of the data word.

33. Media signal according to any one of claims 29 to 32, wherein the plurality of data words comprise secondary media data.

34. Media signal according to any one of claims 29 to 33, wherein the media signal comprises a further stream of the primary media data, wherein the primary media data comprises audio data or video data, wherein the further stream comprising primary media data is aligned to the stream of encoded secondary media data at the predetermined timing instant in the primary media data.

35. Method for decoding a media signal comprising a received stream of digital words representing encoded secondary media data comprising metadata or control data for primary media data and a reference pattern indicating a reference amplitude or a predetermined timing instant in the primary media data, the method comprising:

recovering the secondary media data, wherein the recovering comprises

analyzing the reference pattern to calculate and compensate an amplification or attenuation or a DC offset applied to the stream of digital words, or

converting a sample rate associated with the digital words, according to the predetermined timing instant of the primary media data indicated in the reference pattern, to a predetermined sample rate

to obtain recovered secondary media data; and

deriving a bitstream from the recovered secondary media data.

36. Method for encoding secondary media data comprising metadata or control data for primary media data, the method comprising:

grouping a bitstream of secondary media data to form grouped secondary media data, representing data words;

generating a reference pattern indicating a reference amplitude or a predetermined timing instant in the primary media data;

low-pass filtering the data words and the reference pattern using a pulse shaping filter to obtain data pulses comprising a length of more than one sample of a

predetermined sample rate, wherein an amplitude of a data pulse is weighted according to the data words and the reference pattern, and wherein the pulse shaping filter is configured to add up consecutive weighted data pulses at instants of the predetermined sample rate to obtain the stream of digital words; and

outputting the encoded secondary media data as the stream of digital words.

37. Computer program having a program code for performing the method according to claim 35 or 36, when the computer program runs on a computer.

38. Data processing system comprising:

an encoder in accordance of any one of the claims 1 to 20; and

a decoder in accordance of any one of the claims 21 to 28.

39. The data processing system of claim 38, further comprising:

the signal manipulator for manipulating the stream of digital words to obtain a manipulated stream of digital words,

wherein the decoder is configured to recover a result stream of digital words from the manipulated stream of digital words.

40. The data processing system of claim 39, wherein the signal manipulator is configured to manipulate by amplitude amplification or amplitude attenuation or offset introduction or offset variation or frequency selective attenuation or amplification or resampling, and

wherein the decoder is configured to recover the stream of digital words manipulated by amplitude amplification or amplitude attenuation or offset introduction or offset variation or frequency selective attenuation or amplification or resampling.

41. The data processing system of claim 39 or 40, wherein the signal manipulator is configured to receive a PCM audio channel and to output a PCM audio channel,

wherein the encoder is configured to output a signal transmittable over the PCM audio channel, and

wherein the decoder is configured to receive the received stream from the PCM audio channel.