CLAIMS

1. A decoder system for providing a stereo signal by complex prediction stereo coding, the decoder system comprising:

an upmix stage (206, 207, 210, 211; 306, 307, 308, 309, 312; 406, 407, 408, 409; 1410; 1421; 1433) adapted to generate the stereo signal based on first frequency-domain representations of a downmix signal (M) and a residual signal (D), each of the first frequency-domain representations comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multidimensional space, the upmix stage comprising:

a module (206; 306, 307; 408) for computing a second frequency-domain representation of the downmix signal based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace;

a weighted summer (210, 211; 308, 309; 406, 407) for computing a side signal (S) on the basis of the first and second frequency-domain representations of the downmix signal, the first frequency-domain representation of the residual signal and a complex prediction coefficient (α) encoded in the bit stream signal; and

a sum-and-difference stage (207; 312; 409) for computing the stereo signal on the basis of the first frequency-domain representation of the downmix signal and the side signal,

wherein the upmix stage is further operable in a pass-through mode, in which said downmix and residual signals are supplied to the sum-and-difference stage directly.

2. The decoder system of claim 1, wherein said downmix signal and residual signal are segmented into time frames,

said upmix stage being further adapted to receive, for each time frame, a two-bit data field associated with that frame and to operate, responsive to the value of the data field, in its active mode or its pass-through mode.

3. The decoder system of claim 1, wherein said downmix and residual signal are segmented into time frames,

said upmix stage being further adapted to receive in an MPEG bit stream, for each time frame, an *ms_mask_present* field associated with that frame and to operate, responsive to the value of the *ms_mask_present* field, in its active mode or its pass-through mode.

4. The decoder system of any one of the preceding claims, further comprising:

a dequantization stage (202; 301; 401) arranged upstream of the upmix stage, for providing said first frequency-domain representations of the downmix signal (M) and residual signal (D) based on a bit stream signal.

5. The decoder system of any one of the preceding claims, wherein:

the first spectral components have real values expressed in the first subspace; and

the second spectral components have imaginary values expressed in the second subspace;

optionally, the first spectral components are obtainable by one of the following:

a discrete cosine transform, DCT, or

a modified discrete cosine transform, MDCT,

and, optionally, the second spectral components are obtainable by one of the following:

a discrete sine transform, DST, or

a modified discrete sine transform, MDST.

6. The decoder system of any one of the preceding claims, further comprising at least one temporal noise shaping, TNS, module (204; 303; 403) arranged upstream of the upmix stage;

at least one further TNS module (410) arranged downstream of the upmix stage; and

a selector arrangement (404, 411) for selectively activating either:

(a) said TNS module(s) upstream of the upmix stage, or

(b) said further TNS module(s) downstream of the upmix stage.

7. The decoder of claim 5, wherein:

the downmix signal is partitioned into successive time frames, each associated with a value of the complex prediction coefficient; and

the module for computing a second frequency-domain representation of the downmix signal is adapted to deactivate itself, responsive to the absolute value of the imaginary part of the complex prediction coefficient being smaller than a predetermined tolerance for a time frame, so that it generates no output for that time frame.

8. The decoder of claim 7, wherein the downmix signal time frames are further partitioned into frequency bands, each accompanied by a value of the complex prediction coefficient; and

the module for computing a second frequency-domain representation of the downmix signal is adapted to deactivate itself, responsive to the absolute value of the imaginary part of the complex prediction coefficient being smaller than a predetermined tolerance for a frequency band of a time frame, so that it generates no output for that frequency band.

9. The decoder system of any one of the preceding claims, wherein:

the first spectral components are transform coefficients arranged in one or more time frames of transform coefficients, each block generated by application of a transform to a time segment of a time-domain signal; and

the module for computing a second frequency-domain representation of the downmix signal is adapted to:

derive one or more first intermediate components from at least some of the first spectral components;

form a combination of said one or more first spectral components according to at least a portion of one or more impulse responses to obtain one or more second intermediate components; and

derive said one or more second spectral components from said one or more second intermediate components.

10. The decoder system of claim 9, wherein the portions of the one or more impulse responses are based on frequency response characteristics of the one or more transforms,

wherein optionally the frequency response characteristics of the transform are dependent on characteristics of one or more analysis window functions that were applied with transform to the time segment of the signal.

11. The decoder system of claim 9 or 10, wherein the module for computing a second frequency-domain representation of the downmix signal is adapted to obtain each time frame of second spectral components based on one of the following:

(a) a contemporaneous time frame of first spectral components;

(b) a contemporaneous and a preceding time frame of first spectral components; and

(c) a contemporaneous, a preceding and a subsequent time frame of first spectral components.

12. The decoder system of any one of the preceding claims, wherein the module for computing a second frequency-domain representation of the downmix signal is adapted to compute an approximate second spectral representation comprising approximate second spectral components determined by combination of at least two temporally adjacent and/or frequency-adjacent first spectral components.

13. The decoder system of any one of the preceding claims, said stereo signal being represented in the time domain and the decoder system further comprising:

a switching assembly (203) arranged between said dequantization stage and said upmix stage, operable to function as either:

(a) a pass-through stage, or

(b) a sum-and-difference stage,

thereby enabling switching between directly and jointly coded stereo input signals;

an inverse transform stage (209) adapted to compute a time-domain representation of the stereo signal; and

a selector arrangement (208) arranged upstream of the inverse transform stage, adapted to selectively connect this to either:

(a) a point downstream of the upmix stage, whereby the stereo signal obtained by complex prediction is supplied to the inverse transform stage; or

(b) a point downstream of the switching assembly (203) and upstream of the upmix stage, whereby a stereo signal obtained by direct stereo coding is supplied to the inverse transform stage.

14. An encoder system for encoding a stereo signal using complex prediction as a signal comprising downmix and residual channels and a complex prediction coefficient, the system including:

an estimator (605; 708; 803; 908) for estimating a complex prediction coefficient (α);

a coding stage (601, 602, 603, 604, 606, 607; 701, 702, 703, 704, 706, 707; 801, 802, 804, 805, 806; 901, 902, 903, 904, 906, 907; 1101, 1104, 1105, 1106, 1107) operable:

(a) to transform the stereo signal into a frequency-domain representation of a downmix (M) and a residual (D) signal, in a relationship determined by the value of the complex prediction coefficient; and

(b) to act as a pass-through stage, thereby supplying, directly to the multiplexer, the stereo signal to be encoded.

15. An encoder system according to claim 14, adapted to encode a stereo signal by a bit stream signal by complex prediction stereo coding and further including:

a multiplexer (608; 709; 807; 909; 1111) for receiving output from the coding stage and the estimator and encoding this by said bit stream signal.

16. The encoder system of claim 14 or 15, wherein the estimator is adapted to determine the complex prediction coefficient by minimization over a time interval of a power of the residual signal or an average power of the residual signal.

17. The encoder system of any one of claims 14–16, wherein:

the stereo signal comprises downmix (M) and side (S) channels;

the coding stage is adapted to receive a first frequency-domain representation of the stereo signal, the first frequency-domain representation comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multidimensional space;

the coding stage further comprises:

a module (903; 703; 804; 903) for computing a second frequency-domain representation of said downmix channel based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace; and

a weighted summer (606, 607; 706, 707; 805, 806; 906, 907) for computing a residual signal (D) on the basis of the first and second frequency-domain representations of the downmix channel, the first frequency-domain representation of the side channel and the complex prediction coefficient;

and

the estimator is adapted to receive the downmix and side channels and determine the complex prediction coefficient in order to minimize, over a time interval, the power of the residual signal or to minimize an average power of the residual signal.

18. The encoder system of any one of claims 14–16, wherein:

the coding stage comprises:

a sum-and-difference stage (601; 701; 901) for converting the stereo signal into a jointly coded stereo signal, comprising downmix (M) and side (S) channels;

a transform stage (602, 603; 702, 703; 902, 903) for providing an oversampled frequency-domain representation of the downmix channel and a critically sampled frequency-domain representation of the side channel, wherein the oversampled frequency-domain representation preferably comprises complex spectral components; and

a weighted summer (606, 607; 706, 707; 906, 907) for computing a residual signal (D) on the basis of said oversampled frequency-domain representation of the downmix channel, said critically sampled frequency-domain representation of the side channel and the complex prediction coefficient;

and

the estimator receives the residual signal and determines the complex prediction coefficient in order to minimize the power of the residual signal or to minimize an average power of the residual signal,

wherein, preferably, the transform stage comprises a modified discrete cosine transform, MDCT, stage (606; 706; 906) arranged in parallel with a modified discrete sine transform, MDST, stage (607; 707; 907) for together providing said oversampled frequency-domain representation of the downmix channel.

19. A decoding method for providing a stereo signal by complex prediction stereo coding, the method comprising the steps of:

receiving first frequency-domain representations of a downmix signal (M) and a residual signal (D), each of the first frequency-domain representa-

tions comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multidimensional space;

receiving a control signal; and

responsive to the value of said control signal:

(a) upmixing said downmix and residual signals, using an upmix stage, to obtain said stereo signal, by performing the sub-steps of:

computing a second frequency-domain representation of the downmix signal based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace;

computing a side signal (S) on the basis of the first and second frequency-domain representations of the downmix signal, the first frequency-domain representation of the residual signal and a complex prediction coefficient (α) encoded in the bit stream signal; and

computing said stereo signal by applying a sum-and-difference transform to the first frequency-domain representation of the downmix signal and the side signal,

or

(b) suspending upmixing.

20. The decoding method of claim 19, wherein:

the first spectral components have real values expressed in the first subspace;

the second spectral components have imaginary values expressed in the second subspace;

optionally, the first spectral components are obtainable by one of the following:

a discrete cosine transform, DCT, or

a modified discrete cosine transform, MDCT,

and, optionally, the second spectral components are obtainable by one of the following:

a discrete sine transform, DST, or

a modified discrete sine transform, MDST.

21. The decoding method of claim 20, wherein:

the downmix signal is partitioned into successive time frames, each associated with a value of the complex prediction coefficient; and

said computing a second frequency-domain representation of the downmix signal is suspended responsive to the absolute value of the imaginary part of the complex prediction coefficient being smaller than a predetermined tolerance for a time frame, so that no output is generated for that time frame.

22. The decoding method of claim 21, wherein:

the downmix signal time frames are further partitioned into frequency bands, each accompanied by a value of the complex prediction coefficient; and

said computing a second frequency-domain representation of the downmix signal is suspended responsive to the absolute value of the imaginary part of the complex prediction coefficient being smaller than a predetermined tolerance for a frequency band of a time frame, so that no output is generated for that frequency band.

23. The decoding method of claim 20, wherein:

the first spectral components are transform coefficients arranged in one or more time frames of transform coefficients, each block generated by application of a transform to a time segment of a time-domain signal; and

said computing a second frequency-domain representation of the downmix signal comprises the sub-steps of:

deriving one or more first intermediate components from at least some of the first spectral components;

forming a combination of said one or more first spectral components according to at least a portion of one or more impulse responses to obtain one or more second intermediate components; and

deriving said one or more second spectral components from said one or more second intermediate components.

24. The decoding method of claim 23, wherein the portions of the one or more impulse responses are based on frequency response characteristics of the one or more transforms,

wherein, optionally, the frequency response characteristics of the transform are dependent on characteristics of one or more analysis window functions that were applied with transform to the time segment of the signal.

25. The decoding method of claim 24, wherein said computing a second-frequency-domain representation uses as input, to obtain each time frame of second spectral components, one of the following:

(a) a contemporaneous time frame of first spectral components;

(b) a contemporaneous and a preceding time frame of first spectral components; and

(c) a contemporaneous, a preceding and a subsequent time frame of first spectral components.

26. The decoding method of any one of claims 19–25, wherein said computing a second frequency-domain representation of the downmix signal includes computing an approximate second spectral representation comprising approximate second spectral components determined by combination of at least two temporally adjacent and/or frequency-adjacent first spectral components.

27. The decoding method of any one of claims 19–26, wherein said stereo signal is represented in the time domain and the method further comprises:

omitting, responsive to said bit stream signal being coded by direct stereo coding or joint stereo coding, said step of upmixing; and

inversely transforming said bit stream signal to obtain said stereo signal.

28. The decoding method of claim 27, further comprising the step of:

omitting, responsive to said bit stream signal being coded by direct stereo coding or joint stereo coding, said steps of transforming said time-domain representation of the downmix signal and of computing a side signal; and

inversely transforming the frequency-domain representations of each of the channels encoded by the bit stream signal the to obtain said stereo signal.

29. An encoding method for encoding a stereo signal by a bit stream signal by complex prediction stereo coding, the method comprising the steps of:

determining a complex prediction coefficient (α);

transforming the stereo signal into a first frequency-domain representation of a downmix channel (M) and a residual channel (D), in a relationship determined by the complex prediction coefficient, the first frequency-domain representation comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multi-dimensional space; and

encoding the downmix and residual channels and the complex prediction coefficient as said bit stream.

30. The encoding method of claim 29, wherein the step of determining a complex prediction coefficient is effected with the aim of minimizing a power of the residual signal or an average power of the residual signal over a time interval.

31. The encoding method of claim 29 or 30, further comprising:

defining or recognizing a partition of the stereo signal into time frames;

selecting, for each time segment, whether to encode the stereo signal in this time segment by at least one of the following options:

direct stereo coding,

joint stereo coding, and

complex-prediction stereo coding,

wherein, if direct stereo coding is selected, the stereo signal is transformed into a frequency-domain representation of a left channel (L) and a right channel (R), which are encoded as said bit stream;

and wherein, if joint stereo coding is selected, the stereo signal is transformed into a frequency-domain representation of a downmix (M) and side (S) channel, which are encoded as said bit stream.

32. The encoding method of claim 31, wherein that option is selected that provides a highest quality according to a predefined psychoacoustic model.

33. The encoding method of claim 29 or 30, further comprising the step of defining or recognizing a partition of the stereo signal into time frames,

wherein:

the stereo signal comprises downmix (M) and side (S) channels; and

said step of transforming the stereo signal into a first frequency-domain representation of a downmix channel (M) and a residual channel comprises the sub-steps of:

computing a second frequency-domain representation of the downmix channel based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace; and

forming a residual signal on the basis of the first and second frequency-domain representations of the downmix channel, the first frequency-domain representation of the side channel and the complex prediction coefficient,

and

said step of determining a complex prediction coefficient is effected for one time frame at a time, by minimizing a power of average power of the residual signal in each time frame.

34. The encoding method of any one of claims 29–33, further comprising the steps of:

converting the stereo signal into a jointly coded stereo signal comprising downmix (M) and side (S) channels;

transforming the downmix channel into an oversampled frequency-domain representation, preferably comprising complex spectral components;

transforming the side channel into a critically sampled, preferably real, frequency-domain representation; and

computing a residual signal (D) on the basis of said oversampled frequency-domain representation of the downmix channel, said critically sampled frequency-domain representation of the side channel and the complex prediction coefficient,

wherein said determination of the complex prediction coefficient is effected by feedback control with respect to the residual signal thus computed and in order to minimize its power or its average power.

35. The encoding method of claim 34, wherein:

said transformation of the downmix channel into an oversampled frequency-domain representation is effected by applying a MDCT and a MDST and concatenating outputs from these.

36. A computer-program product comprising a computer-readable medium storing instructions which when executed by a general-purpose computer perform the method set forth in any one of claims 19 to 35.

CLAIMS

1. A decoder system for providing a stereo signal by complex prediction stereo coding, the decoder system comprising:

an upmix stage (206, 207, 210, 211; 306, 307, 308, 309, 312; 406, 407, 408, 409; 1410; 1421; 1433) adapted to generate the stereo signal based on first frequency-domain representations of a downmix signal (M) and a residual signal (D), each of the first frequency-domain representations comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multidimensional space, the upmix stage comprising:

a module (206; 306, 307; 408) for computing a second frequency-domain representation of the downmix signal based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace;

a weighted summer (210, 211; 308, 309; 406, 407) for computing a side signal (S) on the basis of the first and second frequency-domain representations of the downmix signal, the first frequency-domain representation of the residual signal and a complex prediction coefficient (α) encoded in the bit stream signal; and

a sum-and-difference stage (207; 312; 409) for computing the stereo signal on the basis of the first frequency-domain representation of the downmix signal and the side signal,

wherein the upmix stage is further operable in a pass-through mode, in which said downmix and residual signals are supplied to the sum-and-difference stage directly.

2. The decoder system of claim 1, wherein said downmix signal and residual signal are segmented into time frames,

said upmix stage being further adapted to receive, for each time frame, a two-bit data field associated with that frame and to operate, responsive to the value of the data field, in its active mode or its pass-through mode.

3. The decoder system of claim 1, wherein said downmix and residual signal are segmented into time frames,

said upmix stage being further adapted to receive in an MPEG bit stream, for each time frame, an *ms_mask_present* field associated with that frame and to operate, responsive to the value of the *ms_mask_present* field, in its active mode or its pass-through mode.

4. The decoder system of any one of the preceding claims, wherein:

the first spectral components have real values expressed in the first subspace; and

the second spectral components have imaginary values expressed in the second subspace;

optionally, the first spectral components are obtainable by one of the following:

a discrete cosine transform, DCT, or

a modified discrete cosine transform, MDCT,

and, optionally, the second spectral components are obtainable by one of the following:

a discrete sine transform, DST, or

a modified discrete sine transform, MDST.

5. The decoder system of any one of the preceding claims, wherein:

the first spectral components are transform coefficients arranged in one or more time frames of transform coefficients, each block generated by application of a transform to a time segment of a time-domain signal; and

the module for computing a second frequency-domain representation of the downmix signal is adapted to:

derive one or more first intermediate components from at least some of the first spectral components;

form a combination of said one or more first spectral components according to at least a portion of one or more impulse responses to obtain one or more second intermediate components; and

derive said one or more second spectral components from said one or more second intermediate components.

6. The decoder system of any one of the preceding claims, said stereo signal being represented in the time domain and the decoder system further comprising:

a switching assembly (203) arranged between said dequantization stage and said upmix stage, operable to function as either:

(a) a pass-through stage, or

(b) a sum-and-difference stage,

thereby enabling switching between directly and jointly coded stereo input signals;

an inverse transform stage (209) adapted to compute a time-domain representation of the stereo signal; and

a selector arrangement (208) arranged upstream of the inverse transform stage, adapted to selectively connect this to either:

(a) a point downstream of the upmix stage, whereby the stereo signal obtained by complex prediction is supplied to the inverse transform stage; or

(b) a point downstream of the switching assembly (203) and upstream of the upmix stage, whereby a stereo signal obtained by direct stereo coding is supplied to the inverse transform stage.

7. An encoder system for encoding a stereo signal using complex prediction as a signal comprising downmix and residual channels and a complex prediction coefficient, wherein the stereo signal comprises a downmix (M) and a side (S) channel, the system including:

an estimator (605; 708; 803; 908) for estimating a complex prediction coefficient (α);

a coding stage (601, 602, 603, 604, 606, 607; 701, 702, 703, 704, 706, 707; 801, 802, 804, 805, 806; 901, 902, 903, 904, 906, 907; 1101, 1104, 1105, 1106, 1107)

operable:

(a) to transform the stereo signal into a first frequency-domain representation of the downmix (M) channel and the side (S) channel; the first frequency-domain representation comprising first spectral components representing spectral content of the corresponding channel expressed in a first subspace of a multidimensional space; and

(b) to act as a pass-through stage, thereby supplying, directly as output, the stereo signal to be encoded;

and comprising:

a module (903; 804; 903) for computing a second frequency-domain representation of said downmix (M) channel, the second frequency-domain representation comprising second spectral components representing spectral content of the channel expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace; and

a weighted summer (606, 607; 706, 707; 805, 806; 906, 907) for computing a residual (D) channel on the basis of the first and second frequency-domain representations of the downmix (M) channel, the first frequency-domain representation of the side (S) channel and the complex prediction coefficient.

8. The encoder system according to claim 7, adapted to encode the stereo signal by a bit stream signal by complex prediction stereo coding and further including:

a multiplexer (608; 709; 807; 909; 1111) for receiving output from the coding stage and the estimator and encoding this by said bit stream signal.

9. The encoder system of claim 7 or 8, wherein the estimator is adapted to determine the complex prediction coefficient by minimization over a time interval of a power of the residual channel or an average power of the residual channel.

10. The encoder system of any one of claims 7–9, wherein the second frequency-domain representation of said downmix channel is determined based on the first frequency-domain representation thereof.

11. A decoding method for providing a stereo signal by complex prediction stereo coding, the method comprising the steps of:

receiving first frequency-domain representations of a downmix signal (M) and a residual signal (D), each of the first frequency-domain representations comprising first spectral components representing spectral content of the corresponding signal expressed in a first subspace of a multidimensional space;

receiving a control signal; and

responsive to the value of said control signal:

(a) upmixing said downmix and residual signals, using an upmix stage, to obtain said stereo signal, by performing the sub-steps of:

computing a second frequency-domain representation of the downmix signal based on the first frequency-domain representation thereof, the second frequency-domain representation comprising second spectral components representing spectral content of the signal expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace;

computing a side signal (S) on the basis of the first and second frequency-domain representations of the downmix signal, the first frequency-domain representation of the residual signal and a complex prediction coefficient (α) encoded in the bit stream signal; and

computing said stereo signal by applying a sum-and-difference transform to the first frequency-domain representation of the downmix signal and the side signal,

or

(b) suspending upmixing.

12. An encoding method for encoding a stereo signal by a bit stream signal by complex prediction stereo coding, wherein the stereo signal comprises downmix (M) and side (S) channels segmented into time frames, the method comprising, for each time frame, the steps of:

determining a complex prediction coefficient (α); and

selecting between:

(a) transforming the stereo signal into a first frequency-domain representation of the downmix (M) channel and the side (S) channel, the first frequency-domain representation comprising first spectral components representing spectral content of the corresponding channel expressed in a first subspace of a multidimensional space;

computing a second frequency-domain representation of said downmix (M) channel, the second frequency-domain representation comprising second spectral components representing spectral content of the channel expressed in a second subspace of the multidimensional space that includes a portion of the multidimensional space not included in the first subspace; computing a residual (D) channel on the basis of the first and second frequency-domain representations of the downmix (M) channel, the first frequency-domain representation of the side (S) channel and the complex prediction coefficient; and

encoding the downmix (M) and residual (D) channels and the complex prediction coefficient as said bit stream; and

(b) encoding the downmix (M) and side (S) channels as said bit stream.

13. A computer-program product comprising a computer-readable medium storing instructions which when executed by a general-purpose computer perform the method set forth in any one of claims 11 to 12.