Claims

1. Audio decoder configured to decode an audio signal at a first sampling rate from a data stream into which the audio signal is transform coded at a second sampling rate, the first sampling rate being $1/F^{th}$ of the second sampling rate, the audio decoder comprising:

a receiver configured to receive, per frame of length N of the audio signal, N spectral coefficients;

a grabber configured to grab-out for each frame, a low-frequency fraction of length N/F out of the N spectral coefficients;

a spectral-to-time modulator configured to subject, for each frame, the low-frequency fraction to an inverse transform having modulation functions of length (E + 2) · N/F temporally extending over the respective frame and E + 1 previous frames so as to obtain a temporal portion of length (E + 2) · N/F;

a windower configured to window, for each frame, the temporal portion using a synthesis window of length (E +2) · N/F comprising a zero-portion of length 1/4·N/F at a leading end thereof and having a peak within a temporal interval of the synthesis window, the temporal interval succeeding the zero-portion and having length 7/4 · N/F so that the windower obtains a windowed temporal portion of length (E + 2) · N/F; and

a time domain aliasing canceler configured to subject the windowed temporal portion of the frames to an overlap-add process so that a trailing-end fraction of length (E + 1)/(E + 2) of the windowed temporal portion of a current frame overlaps a leading end of length (E + 1)/(E + 2) of the windowed temporal portion of a preceding frame,

wherein the inverse transform is an inverse MDCT or inverse MDST, and

wherein the synthesis window is a downsampled version of a reference synthesis window of length (E + 2) · N, downsampled by a factor of F by a segmental interpolation in segments of length 1/4 · N.

2. Audio decoder according to claim 1, wherein the synthesis window is a concatenation of spline functions of length 1/4 · N/F.

3. Audio decoder according to claim 1 or 2, wherein the synthesis window is a concatenation of cubic spline functions of length 1/4 · N/F.

4. Audio decoder according to any of the previous claims, wherein E = 2.

5. Audio decoder according to any of the previous claims, wherein the inverse transform is an inverse MDCT.

6. Audio decoder according to any of the previous claims, wherein more than 80% of a mass of the synthesis window is comprised within the temporal interval succeeding the zero-portion and having length 7/4 · N/F.

7. Audio decoder according to any of the previous claims, wherein the audio decoder is configured to perform the interpolation or to derive the synthesis window from a storage.

8. Audio decoder according to any of the previous claims, wherein the audio decoder is configured to support different values for F.

9. Audio decoder according to any of the previous claims, wherein F is between 1.5 and 10, both inclusively.

10. Audio decoder according to any of the previous claims, wherein the reference synthesis window is unimodal.

11. Audio decoder according to any of the previous claims, wherein the audio decoder is configured to perform the interpolation in such a manner that a majority of the coefficients of the synthesis window depends on more than two coefficients of the reference synthesis window.

12. Audio decoder according to any of the previous claims, wherein the audio decoder is configured to perform the interpolation in such a manner that each coefficient of

the synthesis window separated by more than two coefficient from segment borders depend on more than two coefficients of the reference synthesis window.

13. Audio decoder according to any of the previous claims, wherein the windower and the time domain aliasing canceller cooperate so that the windower skips the zero-portion in weighting the temporal portion using the synthesis window and the time domain aliasing canceler disregards a corresponding non-weighted portion of the windowed temporal portion in the overlap-add process so that merely E+1 windowed temporal portions are summed-up so as to result in the corresponding non-weighted portion of a corresponding frame and E+2 windowed portions are summed-up within a reminder of the corresponding frame.

14. Audio decoder for generating a downscaled version of a synthesis window of an audio decoder according to any of the previous claims, wherein E=2 so that the synthesis window function comprises a kernel related half of length $2 \cdot N/F$ preceded by a reminder half of length $2 \cdot N/F$ and wherein the spectral-to-time modulator, the windower and the time domain aliasing canceler are implemented so as to cooperate in a lifting implementation according to which

the spectral-to-time modulator confines the subjecting, for each frame, the low-frequency fraction to the inverse transform having modulation functions of length $(E + 2) \cdot N/F$ temporally extending over the respective frame and $E + 1$ previous frames, to a transform kernel coinciding with the respective frame and one previous frame so as to obtain the temporal portion $x_{k,n}$ with $n = 0 \ldots 2M-1$ with $M=N/F$ being a sample index and k being a frame index;

the windower windowing, for each frame, the temporal portion $x_{k,n}$ according to $z_{k,n} = \omega_n \cdot x_{k,n}$ for $n = 0,\ldots,2M-1$ so as to obtain the windowed temporal portion $z_{k,n}$ with with $n = 0 \ldots 2M-1$;

the time domain aliasing canceler generates intermediate temporal portions $m_k(0),\ldots m_k(M-1)$ according to $m_{k,n} = z_{k,n} + z_{k-1,n+M}$ for $n = 0,\ldots,M-1$, and

the audio decoder comprises a lifter configured to obtain the frames $u_{k,n}$ with $n = 0 \ldots M-1$ according to

$$u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n} \qquad \text{for } n = M/2,\ldots,M-1 \quad ,$$

and

$$u_{k,n} = m_{k,n} + l_{M-1-n} \cdot m_{k-1,M-1-n} \qquad \text{for } n=0,\ldots,M/2-1 \quad ,$$

wherein l_n with n = 0…M-1 are lifting coefficients, and wherein l_n with n = 0…M-1 and ω_n with n = 0,…,2M-1 depend on coefficients w_n with n = 0…(E+2)M-1 of the synthesis window.

15. Audio decoder configured to decode an audio signal at a first sampling rate from a data stream into which the audio signal is transform coded at a second sampling rate, the first sampling rate being $1/F^{th}$ of the second sampling rate, the audio decoder comprising:

a receiver configured to receive, per frame of length N of the audio signal, N spectral coefficients;

a grabber configured to grab-out for each frame, a low-frequency fraction of length N/F out of the N spectral coefficients;

a spectral-to-time modulator configured to subject, for each frame, the low-frequency fraction to an inverse transform having modulation functions of length 2·N/F temporally extending over the respective frame and a previous frame so as to obtain a temporal portion of length 2·N/F;

a windower configured to window, for each frame, the temporal portion $x_{k,n}$ according to $z_{k,n} = \omega_n \cdot x_{k,n}$ for n = 0,…,2M-1 so as to obtain a windowed temporal portion $z_{k,n}$ with with n = 0…2M-1;

a time domain aliasing canceler configured to generate intermediate temporal portions $m_k(0),\ldots m_k(M-1)$ according to $m_{k,n} = z_{k,n} + z_{k-1,n+M}$ for n = 0,…,M-1, and

a lifter configured to obtain frames $u_{k,n}$ of the audio signal with n = 0…M-1 according to

$$u_{k,n} = m_{k,n} + l_{n-M/2} \cdot m_{k-1,M-1-n} \quad \text{for } n = M/2,\ldots,M-1 \quad ,$$

and

$$u_{k,n} = m_{k,n} + l_{M-1-n} \cdot m_{k-1,M-1-n} \quad \text{for } n=0,\ldots,M/2-1 \quad ,$$

wherein l_n with n = 0…M-1 are lifting coefficients,

wherein the inverse transform is an inverse MDCT or inverse MDST, and

wherein l_n with n = 0…M-1 and ω_n with n = 0,…,2M-1 depend on coefficients w_n with n = 0…(E+2)M-1 of a synthesis window, with E being an integer, and the synthesis window is a downsampled version of a reference synthesis window of length 4 · N, downsampled by a factor of F by a segmental interpolation in segments of length 1/4 · N.

16. Apparatus for generating a downscaled version of a synthesis window of an audio decoder according to any of the previous claims, wherein the synthesis window of length (E + 2) · N and the apparatus is configured to downsample the synthesis window by a factor of F by a segmental interpolation in 4 · (E + 2) segments of equal length.

17. Method for generating a downscaled version of a synthesis window of an audio decoder according to any of claims 1 to 15, wherein the synthesis window of length (E + 2) · N and the method comprises downsampling the synthesis window by a factor of F by a segmental interpolation in 4 · (E + 2) segments of equal length.

18. Method for decoding an audio signal at a first sampling rate from a data stream into which the audio signal is transform coded at a second sampling rate, the first sampling rate being $1/F^{th}$ of the second sampling rate, the method comprising:

receiving, per frame of length N of the audio signal, N spectral coefficients;

grabbing-out for each frame, a low-frequency fraction of length N/F out of the N spectral coefficients;

performing a spectral-to-time modulation by subjecting, for each frame, the low-frequency fraction to an inverse transform having modulation functions of length (E + 2) · N/F temporally extending over the respective frame and E + 1 previous frames so as to obtain a temporal portion of length (E + 2) · N/F;

windowing, for each frame, the temporal portion using a synthesis window of length (E +2) · N/F comprising a zero-portion of length 1/4·N/F at a leading end thereof and having a peak within a temporal interval of the synthesis window, the temporal interval succeeding the zero-portion and having length 7/4 · N/F so that the windower obtains a windowed temporal portion of length (E + 2) · N/F; and

performing a time domain aliasing cancellation by subjecting the windowed temporal portion of the frames to an overlap-add process so that a trailing-end fraction of length (E + 1)/(E + 2) of the windowed temporal portion of a current frame overlaps a leading end of length (E + 1)/(E + 2) of the windowed temporal portion of a preceding frame,

wherein the inverse transform is an inverse MDCT or inverse MDST, and

wherein the synthesis window is a downsampled version of a reference synthesis window of length (E + 2) · N, downsampled by a factor of F by a segmental interpolation in segments of length 1/4 · N.

19. Computer program having a program code for performing, when running on a computer, a method according to claim 17 or 18.