CLAIMS

1. A method for performing high frequency reconstruction of an audio signal, the method comprising:

receiving an encoded audio bitstream, the encoded audio bitstream including audio data representing a lowband portion of the audio signal and high frequency reconstruction metadata;

decoding the audio data to generate a decoded lowband audio signal;

extracting from the encoded audio bitstream the high frequency reconstruction metadata, the high frequency reconstruction metadata including operating parameters for a high frequency reconstruction process, the operating parameters including a patching mode parameter located in a backward-compatible extension container of the encoded audio bitstream, wherein a first value of the patching mode parameter indicates spectral translation and a second value of the patching mode parameter indicates harmonic transposition by phase-vocoder frequency spreading;

filtering the decoded lowband audio signal to generate a filtered lowband audio signal;

regenerating a highband portion of the audio signal using the filtered lowband audio signal and the high frequency reconstruction metadata, wherein the regenerating includes spectral translation if the patching mode parameter is the first value and the regenerating includes harmonic transposition by phase-vocoder frequency spreading if the patching mode parameter is the second value; and

combining the filtered lowband audio signal with the regenerated highband portion to form a wideband audio signal,

wherein the filtering, regenerating, and combining are performed as a post-processing operation with a delay of 3010 samples per audio channel.

2. The method of claim 1 wherein the encoded audio bitstream further includes a fill element with an identifier indicating a start of the fill element and fill data after the identifier, wherein the fill data includes the backward-compatible extension container.

3. The method of claim 2 wherein the identifier is a three bit unsigned integer transmitted most significant bit first and having a value of 0x6.

4. The method of claim 2 or claim 3, wherein the fill data includes an extension payload, the extension payload includes spectral band replication extension data, and the extension payload is identified with a four bit unsigned integer transmitted most significant bit first and having a value of ‘1101’ or ‘1110’, and, optionally,

wherein the spectral band replication extension data includes:

an optional spectral band replication header,

spectral band replication data after the header, and

a spectral band replication extension element after the spectral band replication data, and wherein the flag is included in the spectral band replication extension element.

5. The method of any one of claims 1-4 wherein the high frequency reconstruction metadata includes envelope scale factors, noise floor scale factors, time/frequency grid information, or a parameter indicating a crossover frequency.

6. The method of any one of claims 1-5 wherein the backward-compatible extension container further includes a flag indicating whether additional preprocessing is used to avoid discontinuities in a shape of a spectral envelope of the highband portion when the patching mode parameter equals the first value, wherein a first value of the flag enables the additional preprocessing and a second value of the flag disables the additional preprocessing.

7. The method of claim 6 wherein the additional preprocessing includes calculating a pre-gain curve using a linear prediction filter coefficient.

8. The method of any one of claims 1-5 wherein the backward-compatible extension container further includes a flag indicating whether signal adaptive frequency domain oversampling is to be applied when the patching mode parameter equals the second value, wherein a first value of the flag enables the signal adaptive frequency domain oversampling and a second value of the flag disables the signal adaptive frequency domain oversampling.

9. The method of claim 8 wherein the signal adaptive frequency domain oversampling is applied only for frames containing a transient.

10. The method of any one of the previous claims wherein the harmonic transposition by phase-vocoder frequency spreading is performed with an estimated complexity at or below 4.5 million of operations per second and 3 kWords of memory.

11. The method of any one of the previous claims wherein

filtering the decoded lowband audio signal to generate a filtered lowband audio signal comprises filtering the decoded lowband audio signal into a plurality of subbands using a complex QMF analysis filter bank; and

combining the filtered lowband audio signal with the regenerated highband portion to form a wideband audio signal comprises using a complex QMF synthesis filter bank.

12. The method of claim 11, wherein analysis filters $h_k(n)$ of the complex QMF analysis filter bank and synthesis filters $f_k(n)$ of the complex QMF synthesis filter bank are defined by:

$$h_k(n) = f_k(n) = p_0(n) \exp\left\{i\frac{\pi}{M}\left(k+\frac{1}{2}\right)\left(n-\frac{N}{2}\right)\right\}, 0 \le n \le N;\ 0 \le k < M$$

where $p_0(n)$ is a real-valued prototype filter, M denotes a number of channels and N is a prototype filter order.

13. A computer program product having instructions which, when executed by a computing device or system, cause said computing device or system to execute the method of any of the claims 1-12.

14. An audio processing unit for performing high frequency reconstruction of an audio signal, the audio processing unit comprising:

an input interface for receiving an encoded audio bitstream, the encoded audio bitstream including audio data representing a lowband portion of the audio signal and high frequency reconstruction metadata;

a core audio decoder for decoding the audio data to generate a decoded lowband audio signal;

a deformatter for extracting from the encoded audio bitstream the high frequency reconstruction metadata, the high frequency reconstruction metadata including operating parameters for a high frequency reconstruction process, the operating parameters including a patching mode parameter located in a backward-compatible extension container of the encoded audio bitstream, wherein a first value of the patching mode parameter indicates spectral translation and a second value of the patching mode parameter indicates harmonic transposition by phase-vocoder frequency spreading;

an analysis filterbank for filtering the decoded lowband audio signal to generate a filtered lowband audio signal;

a high frequency regenerator for reconstructing a highband portion of the audio signal using the filtered lowband audio signal and the high frequency reconstruction metadata, wherein the reconstructing includes a spectral translation if the patching mode parameter is the first value and the reconstructing includes harmonic transposition by phase-vocoder frequency spreading if the patching mode parameter is the second value; and

a synthesis filterbank for combining the filtered lowband audio signal with the regenerated highband portion to form a wideband audio signal,

wherein the analysis filterbank, high frequency regenerator, and synthesis filterbank are performed in a post-processor with a delay of 3010 samples per audio channel.

15. The audio processing unit of claim 14 wherein the harmonic transposition by phase-vocoder frequency spreading is performed with an estimated complexity at or below 4.5 million of operations per second and 3 kWords of memory.

CLAIMS

1. A method for performing high frequency reconstruction of an audio signal, the method comprising:

receiving an encoded audio bitstream, the encoded audio bitstream including audio data representing a lowband portion of the audio signal and high frequency reconstruction metadata;

decoding the audio data to generate a decoded lowband audio signal;

extracting from the encoded audio bitstream the high frequency reconstruction metadata, the high frequency reconstruction metadata including operating parameters for a high frequency reconstruction process, the operating parameters including a patching mode parameter located in a backward-compatible extension container of the encoded audio bitstream, wherein a first value of the patching mode parameter indicates spectral translation and a second value of the patching mode parameter indicates harmonic transposition by phase-vocoder frequency spreading;

filtering the decoded lowband audio signal to generate a filtered lowband audio signal;

regenerating a highband portion of the audio signal using the filtered lowband audio signal and the high frequency reconstruction metadata, wherein the regenerating includes spectral translation if the patching mode parameter is the first value and the regenerating includes harmonic transposition by phase-vocoder frequency spreading if the patching mode parameter is the second value; and

combining the filtered lowband audio signal with the regenerated highband portion to form a wideband audio signal,

wherein the filtering, regenerating, and combining are performed as a post-processing operation with a delay of 3010 samples per audio channel.

2. The method of claim 1 wherein the encoded audio bitstream further includes a fill element with an identifier indicating a start of the fill element and fill data after the identifier, wherein the fill data includes the backward-compatible extension container.

3. The method of claim 1 wherein the harmonic transposition by phase-vocoder frequency spreading is performed with an estimated complexity at or below 4.5 million of operations per second and at or below 3 kWords of memory.

4. A non-transitory computer-readable medium having instructions which, when executed by a computing device or system, cause said computing device or system to execute the method of claim 1.

5. An audio processing unit for performing high frequency reconstruction of an audio signal, the audio processing unit comprising:

an input interface for receiving an encoded audio bitstream, the encoded audio bitstream including audio data representing a lowband portion of the audio signal and high frequency reconstruction metadata;

a core audio decoder for decoding the audio data to generate a decoded lowband audio signal;

a deformatter for extracting from the encoded audio bitstream the high frequency reconstruction metadata, the high frequency reconstruction metadata including operating parameters for a high frequency reconstruction process, the operating parameters including a patching mode parameter located in a backward-compatible extension container of the encoded audio bitstream, wherein a first value of the patching mode parameter indicates spectral translation and a second value of the patching mode parameter indicates harmonic transposition by phase-vocoder frequency spreading;

an analysis filterbank for filtering the decoded lowband audio signal to generate a filtered lowband audio signal;

a high frequency regenerator for reconstructing a highband portion of the audio signal using the filtered lowband audio signal and the high frequency reconstruction metadata, wherein the reconstructing includes a spectral translation if the patching mode parameter is the first value and the reconstructing includes harmonic transposition by phase-vocoder frequency spreading if the patching mode parameter is the second value; and

a synthesis filterbank for combining the filtered lowband audio signal with the regenerated highband portion to form a wideband audio signal,

wherein the analysis filterbank, high frequency regenerator, and synthesis filterbank are performed in a post-processor with a delay of 3010 samples per audio channel, so that a composition time applies to a 3011-th audio sample within an audio composition unit.

6. The audio processing unit of claim 5 wherein the harmonic transposition by phase-vocoder frequency spreading is performed with an estimated complexity at or below 4.5 million of operations per second and at or below 3 kWords of memory.