

FREQUENCY-BASED CODING OF AUDIO CHANNELS IN PARAMETRIC MULTI-CHANNEL CODING SYSTEMS

BACKGROUND OF THE INVENTION

Field of the Invention

5 The present invention relates to the encoding of audio signals and the subsequent synthesis of auditory scenes from the encoded audio data.

Cross-Reference to Related Applications

 This application claims the benefit of the filing date of U.S. provisional application no. 60/549,972, filed on 03/04/04 as attorney docket no. Faller 14-2. The subject matter of this application
10 is related to the subject matter of U.S. patent application serial number 09/848,877, filed on 05/04/2001 as attorney docket no. Faller 5 ("the '877 application"), U.S. patent application serial number 10/045,458, filed on 11/07/2001 as attorney docket no. Baumgarte 1-6-8 ("the '458 application"), and U.S. patent application serial number 10/155,437, filed on 05/24/2002 as attorney docket no. Baumgarte 2-10 ("the '437 application"), and U.S. patent application serial number 10/815,591, filed on 04/01/2004
15 as attorney docket no. Baumgarte 7-12 ("the '591 application"), the teachings of all four of which are incorporated herein by reference.

Description of the Related Art

 Multi-channel surround audio systems have been standard in movie theaters for years. As technology has advanced, it has become affordable to produce multi-channel surround systems for home
20 use. Today, such systems are mostly sold as "home theater systems." Conforming to an ITU-R recommendation, the vast majority of these systems provide five regular audio channels and one low-frequency sub-woofer channel (denoted the low-frequency effects or LFE channel). Such multi-channel system is denoted a 5.1 surround system. There are other surround systems, such as 7.1 (seven regular channels and one LFE channel) and 10.2 (ten regular channels and two LFE channels).

25 C. Faller and F. Baumgarte, "Efficient representation of spatial audio coding using perceptual parametrization," *IEEE Workshop on Appl. of Sig. Proc. to Audio and Acoust.*, October 2001, and C. Faller and F. Baumgarte, "Binaural Cue Coding Applied to Stereo and Multi-Channel Audio Compression," *Preprint 112th Conv. Aud. Eng. Soc.*, May 2002, (collectively, "the BCC papers") the teachings of both of which are incorporated herein by reference, describe a parametric multi-channel
30 audio coding technique (referred to as BCC coding).

 Fig. 1 shows a block diagram of an audio processing system **100** that performs binaural cue coding (BCC) according to the BCC papers. BCC system **100** has a BCC encoder **102** that receives C

-2-

audio input channels **108**, for example, one from each of *C* different microphones **106**. BCC encoder **102** has a downmixer **110**, which converts the *C* audio input channels into a mono audio sum signal **112**.

In addition, BCC encoder **102** has a BCC analyzer **114**, which generates BCC cue code data stream **116** for the *C* input channels. The BCC cue codes (also referred to as auditory scene parameters) include inter-channel level difference (ICLD) and inter-channel time difference (ICTD) data for each input channel. BCC analyzer **114** performs band-based processing to generate ICLD and ICTD data for each of one or more different frequency sub-bands (e.g., different critical bands) of the audio input channels.

BCC encoder **102** transmits sum signal **112** and the BCC cue code data stream **116** (e.g., as either in-band or out-of-band side information with respect to the sum signal) to a BCC decoder **104** of BCC system **100**. BCC decoder **104** has a side-information processor **118**, which processes data stream **116** to recover the BCC cue codes **120** (e.g., ICLD and ICTD data). BCC decoder **104** also has a BCC synthesizer **122**, which uses the recovered BCC cue codes **120** to synthesize *C* audio output channels **124** from sum signal **112** for rendering by *C* loudspeakers **126**, respectively.

Audio processing system **100** can be implemented in the context of multi-channel audio signals, such as 5.1 surround sound. In particular, downmixer **110** of BCC encoder **102** would convert the six input channels of conventional 5.1 surround sound (i.e., five regular channels + one LFE channel) into sum signal **112**. In addition, BCC analyzer **114** of encoder **102** would transform the six input channels into the frequency domain to generate the corresponding BCC cue codes **116**. Analogously, side-information processor **118** of BCC decoder **104** would recover the BCC cue codes **120** from the received side information stream **116**, and BCC synthesizer **122** of decoder **104** would (1) transform the received sum signal **112** into the frequency domain, (2) apply the recovered BCC cue codes **120** to the sum signal in the frequency domain to generate six frequency-domain signals, and (3) transform those frequency-domain signals into six time-domain channels of synthesized 5.1 surround sound (i.e., five synthesized regular channels + one synthesized LFE channel) for rendering by loudspeakers **126**.

SUMMARY OF THE INVENTION

For surround sound applications, embodiments of the present invention involve a BCC-based parametric audio coding technique in which band-based BCC coding is not applied to low-frequency sub-woofer (LFE) channel(s) for frequency sub-bands above a cut-off frequency. For example, for 5.1 surround sound, BCC coding is applied to all six channels (i.e., the five regular channels plus the one LFE channel) for sub-bands below the cut-off frequency, while BCC coding is applied to only the five regular channels (i.e., and not to the LFE channel) for sub-bands above the cut-off frequency. By avoiding BCC coding of the LFE channel at "high" frequencies, these embodiments of the present

invention have (1) reduced processing loads at both the encoder and decoder and (2) smaller BCC code bitstreams than corresponding BCC-based systems that process all six channels at all frequencies.

More generally, the present invention involves the application of parametric audio coding techniques, such as BCC coding, but not necessarily limited to BCC coding, where two or more different subsets of input channels are processed for two or more different frequency ranges. As used in this specification, the term "subset" may refer to the set containing all of the input channels as well as to those proper subsets that include fewer than all of the input channels. The application of the present invention to BCC coding of 5.1 and other surround sound signals is just one particular example of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Other aspects, features, and advantages of the present invention will become more fully apparent from the following detailed description, the appended claims, and the accompanying drawings in which:

Fig. 1 shows a block diagram of an audio processing system that performs binaural cue coding (BCC); and

Fig. 2 shows a block diagram of an audio processing system that performs BCC coding according to one embodiment of the present invention.

DETAILED DESCRIPTION

Fig. 2 shows a block diagram of an audio processing system **200** that performs binaural cue coding (BCC) for 5.1 surround audio, according to one embodiment of the present invention. BCC system **200** has a BCC encoder **202**, which receives six audio input channels **208** (i.e., five regular channels and one LFE channel). BCC encoder **202** has a downmixer **210**, which converts (e.g., averages) the audio input channels (including the LFE channel) into one or more, but fewer than six, combined channels **212**.

In addition, BCC encoder **202** has a BCC analyzer **214**, which generates BCC cue code data stream **216** for the input channels. As indicated in Fig. 2, for frequency sub-bands at or below a specified cut-off frequency f_c , BCC analyzer **214** uses all six 5.1 surround sound input channels (including the LFE channel) when generating the BCC cue code data. For all other (i.e., high-frequency) sub-bands, BCC analyzer **214** uses only the five regular channels (and not the LFE channel) to generate the BCC cue code data. As a result, the LFE channel contributes BCC codes for only BCC sub-bands at or below the cut-off-frequency rather than for the full BCC frequency range, thereby reducing the overall size of the side-information bitstream.

The cut-off frequency is preferably chosen such that the effective audio bandwidth of the LFE channel is smaller than or equal to f_c (that is, the LFE channel has substantially zero energy or insubstantial audio content beyond the cut-off frequency). Unless the frequency sub-bands are aligned with the cut-off frequency, the cut-off frequency falls within a particular frequency sub-band. In that

case, part of that sub-band will exceed the cut-off frequency. For purposes of this specification, such a sub-band is referred to as being "at" the cut-off frequency. In preferred embodiments, that entire sub-band of the LFE channel is BCC coded, and the next higher frequency sub-band is the first high-frequency sub-band that is not BCC coded.

5 In one possible implementation, the BCC cue codes include inter-channel level difference (ICLD), inter-channel time difference (ICTD), and inter-channel correlation (ICC) data for the input channels. BCC analyzer **214** preferably performs band-based processing analogous to that described in the '877 and '458 applications to generate ICLD and ICTD data for different frequency sub-bands of the audio input channels. In addition, BCC analyzer **214** preferably generates coherence measures as the
10 ICC data for the different frequency sub-bands. These coherence measures are described in greater detail in the '437 and '591 applications.

BCC encoder **202** transmits the one or more combined channels **212** and the BCC cue code data stream **216** (e.g., as either in-band or out-of-band side information with respect to the combined channels) to a BCC decoder **204** of BCC system **200**. BCC decoder **204** has a side-information
15 processor **218**, which processes data stream **216** to recover the BCC cue codes **220** (e.g., ICLD, ICTD, and ICC data). BCC decoder **204** also has a BCC synthesizer **222**, which uses the recovered BCC cue codes **220** to synthesize six audio output channels **224** from the one or more combined channels **212** for rendering by six surround-sound loudspeakers **226**, respectively.

As indicated in Fig. 2, BCC synthesizer **222** performs six-channel BCC synthesis for sub-bands
20 at or below the cut-off frequency f_c to generate frequency content for all six 5.1 surround channels (i.e., including the LFE channel), while performing five-channel BCC synthesis for sub-bands above the cut-off frequency to generate frequency content for only the five regular channels of 5.1 surround sound. In particular, BCC synthesizer **222** decomposes the received combined channel(s) **212** into a number of frequency sub-bands (e.g., critical bands). In these sub-bands, different processing is applied to obtain
25 the corresponding sub-bands of the output audio channels. The result is that, for the LFE channel, only sub-bands with frequencies at or below the cut-off frequency are obtained. In other words, the LFE channel has frequency content only for sub-bands at or below the cut-off frequency. The upper sub-bands of the LFE channel (i.e., those above the cut-off frequency) may be filled with zero signals (if necessary).

30 Depending on the particular implementation, a BCC encoder could be designed to generate BCC cue codes for all frequencies and simply not transmit those codes for particular sub-bands (e.g., sub-bands above the cut-off frequency and/or sub-bands having substantially zero energy). Similarly, the corresponding BCC decoder could be designed to perform conventional BCC synthesis for all frequencies, where the BCC decoder applies appropriate BCC cue code values for those sub-bands having no
35 explicitly transmitted codes.

Although the present invention has been described in the context of BCC decoders that apply the techniques of the '877 and '458 applications to synthesize auditory scenes, the present invention can also be implemented in the context of BCC decoders that apply other techniques for synthesizing auditory scenes that do not necessarily rely on the techniques of the '877 and '458 applications. For example, the BCC processing of the present invention can be implemented without ICTD, ICLD, and/or ICC data, with or without other suitable cue codes, such as, for example, those associated with head-related transfer functions.

In the embodiment of Fig. 2, 5.1 surround sound is encoded by applying six-channel BCC analysis to sub-bands at or below the cut-off frequency and five-channel BCC analysis to sub-bands above the cut-off frequency. In another embodiment, the present invention can be applied to 7.1 surround sound in which eight-channel BCC analysis is applied to sub-bands at or below a specified cut-off frequency and seven-channel BCC analysis (excluding the single LFE channel) is applied to sub-bands above the cut-off frequency.

The present invention can also be applied to surround audio having more than one LFE channel. For example, for 10.2 surround sound, twelve-channel BCC analysis could be applied to sub-bands at or below a specified cut-off frequency, while ten-channel BCC analysis (excluding the two LFE channels) could be applied to sub-bands above the cut-off frequency. Alternatively, there could be two different cut-off frequencies specified: a first cut-off frequency for a first LFE channel of the 10.2 surround sound and second cut-off frequency for the second LFE channel. In this case and assuming that the first cut-off frequency is lower than the second cut-off frequency, twelve-channel BCC analysis could be applied to sub-bands at or below the first cut-off frequency, eleven-channel BCC analysis (excluding the first LFE channel) could be applied to sub-bands that are (1) above the first cut-off frequency and (2) at or below the second cut-off frequency, and ten-channel BCC analysis (excluding both LFE channels) could be applied to sub-bands above the second cut-off frequency.

Similarly, some consumer multi-channel equipment is purposely designed with different output channels having different frequency ranges. For example, some 5.1 surround sound equipment have two rear channels that are designed to reproduce only frequencies below 7kHz. The present invention could be applied to such systems by specifying two cut-off frequencies: one for the LFE channel and a higher one for the rear channels. In this case, six-channel BCC analysis could be applied to sub-bands at or below the LFE cut-off frequency, five-channel BCC analysis (excluding the LFE channel) could be applied to sub-bands that are (1) above the LFE cut-off frequency and (2) at or below the rear-channel cut-off frequency, and three-channel BCC analysis (excluding the LFE channel and the two rear channels) could be applied to sub-bands above the rear-channel cut-off frequency.

The present invention can be generalized further to apply parametric audio coding to two or more different subsets of input channels for two or more different frequency regions, where the

parametric audio coding could be other than BCC coding and the different frequency regions are chosen such that the frequency content of the different input channels is reflected in these regions. Depending on the particular application, different channels could be excluded from different frequency regions in any suitable combinations. For example, low-frequency channels could be excluded from high-frequency regions and/or high-frequency channels could be excluded from low-frequency regions. It may even be the case that no single frequency region involves all of the input channels.

As described previously, although the input channels **208** can be downmixed to form a single combined (e.g., mono) channel **212**, in alternative implementations, the multiple input channels can be downmixed to form two or more different "combined" channels, depending on the particular audio processing application. More information on such techniques can be found in U.S. patent application no. 10/762,100, filed on 01/20/04, the teachings of which are incorporated herein by reference.

In some implementations, when downmixing generates multiple combined channels, the combined channel data can be transmitted using conventional audio transmission techniques. For example, when two combined channels are generated, conventional stereo transmission techniques may be able to be employed. In this case, a BCC decoder can extract and use the BCC codes to synthesize a multi-channel signal (e.g., 5.1 surround sound) from the two combined channels. Moreover, this can provide backwards compatibility, where the two BCC combined channels are played back using conventional (i.e., non-BCC-based) stereo decoders that ignore the BCC codes. Analogously, backwards compatibility can be achieved for a conventional mono decoder when a single BCC combined channel is generated. Note that, in theory, when there are multiple "combined" channels, one or more of the combined channels may actually be based on individual input channels.

Although BCC system **200** can have the same number of audio input channels as audio output channels, in alternative embodiments, the number of input channels could be either greater than or less than the number of output channels, depending on the particular application. For example, the input audio could correspond to 7.1 surround sound and the synthesized output audio could correspond to 5.1 surround sound, or vice versa.

In general, BCC encoders of the present invention may be implemented in the context of converting M input audio channels into N combined audio channels and one or more corresponding sets of BCC codes, where $M > N \geq 1$. Similarly, BCC decoders of the present invention may be implemented in the context of generating P output audio channels from the N combined audio channels and the corresponding sets of BCC codes, where $P > N$, and P may be the same as or different from M .

Depending on the particular implementation, the various signals received and generated by both BCC encoder **202** and BCC decoder **204** of Fig. 2 may be any suitable combination of analog and/or digital signals, including all analog or all digital. Although not shown in Fig. 2, those skilled in the art will appreciate that the one or more combined channels **212** and the BCC cue code data stream **216** may

be further encoded by BCC encoder **202** and correspondingly decoded by BCC decoder **204**, for example, based on some appropriate compression scheme (e.g., ADPCM) to further reduce the size of the transmitted data.

The definition of transmission of data from BCC encoder **202** to BCC decoder **204** will depend on the particular application of audio processing system **200**. For example, in some applications, such as live broadcasts of music concerts, transmission may involve real-time transmission of the data for immediate playback at a remote location. In other applications, "transmission" may involve storage of the data onto CDs or other suitable storage media for subsequent (i.e., non-real-time) playback. Of course, other applications may also be possible.

Depending on the particular implementation, the transmission channels may be wired or wireless and can use customized or standardized protocols (e.g., IP). Media like CD, DVD, digital tape recorders, and solid-state memories can be used for storage. In addition, transmission and/or storage may, but need not, include channel coding. Similarly, although the present invention has been described in the context of digital audio systems, those skilled in the art will understand that the present invention can also be implemented in the context of analog audio systems, such as AM radio, FM radio, and the audio portion of analog television broadcasting, each of which supports the inclusion of an additional in-band low-bitrate transmission channel.

The present invention can be implemented for many different applications, such as music reproduction, broadcasting, and telephony. For example, the present invention can be implemented for digital radio/TV/internet (e.g., Webcast) broadcasting such as Sirius Satellite Radio or XM. Other applications include voice over IP, PSTN or other voice networks, analog radio broadcasting, and Internet radio.

Depending on the particular application, different techniques can be employed to embed the sets of BCC codes into a combined channel to achieve a BCC signal of the present invention. The availability of any particular technique may depend, at least in part, on the particular transmission/storage medium(s) used for the BCC signal. For example, the protocols for digital radio broadcasting usually support inclusion of additional enhancement bits (e.g., in the header portion of data packets) that are ignored by conventional receivers. These additional bits can be used to represent the sets of auditory scene parameters to provide a BCC signal. In general, the present invention can be implemented using any suitable technique for watermarking of audio signals in which data corresponding to the sets of auditory scene parameters are embedded into the audio signal to form a BCC signal. For example, these techniques can involve data hiding under perceptual masking curves or data hiding in pseudo-random noise. The pseudo-random noise can be perceived as comfort noise. Data embedding can also be implemented using methods similar to bit robbing used in TDM (time division

-8-

multiplexing) transmission for in-band signaling. Another possible technique is mu-law LSB bit flipping, where the least significant bits are used to transmit data.

The present invention may be implemented as circuit-based processes, including possible implementation on a single integrated circuit. As would be apparent to one skilled in the art, various functions of circuit elements may also be implemented as processing steps in a software program. Such software may be employed in, for example, a digital signal processor, micro-controller, or general-purpose computer.

The present invention can be embodied in the form of methods and apparatuses for practicing those methods. The present invention can also be embodied in the form of program code embodied in tangible media, such as floppy diskettes, CD-ROMs, hard drives, or any other machine-readable storage medium, wherein, when the program code is loaded into and executed by a machine, such as a computer, the machine becomes an apparatus for practicing the invention. The present invention can also be embodied in the form of program code, for example, whether stored in a storage medium, loaded into and/or executed by a machine, or transmitted over some transmission medium or carrier, such as over electrical wiring or cabling, through fiber optics, or via electromagnetic radiation, wherein, when the program code is loaded into and executed by a machine, such as a computer, the machine becomes an apparatus for practicing the invention. When implemented on a general-purpose processor, the program code segments combine with the processor to provide a unique device that operates analogously to specific logic circuits.

It will be further understood that various changes in the details, materials, and arrangements of the parts which have been described and illustrated in order to explain the nature of this invention may be made by those skilled in the art without departing from the scope of the invention as expressed in the following claims.

CLAIMS

What is claimed is:

1. A method for encoding a multi-channel audio signal having a plurality of audio input channels, the method comprising:

5 applying a parametric audio encoding technique to generate parametric audio codes for a first subset of the audio input channels for a first frequency region; and

 applying the parametric audio encoding technique to generate parametric audio codes for a second subset of the audio input channels for a second frequency region, wherein:

 the second frequency region is different from the first frequency region; and

10 the second subset is different from the first subset.

2. The invention of claim 1, wherein the parametric audio encoding technique is binaural cue coding (BCC) encoding.

3. The invention of claim 1, wherein:

15 the multi-channel audio signal is a surround sound signal having a plurality of regular channels and at least one low-frequency (LFE) channel;

 the first subset includes all of the audio input channels;

 the first frequency region corresponds to sub-bands at or below a specified cut-off frequency;

 the second subset excludes the LFE channel; and

 the second frequency region corresponds to sub-bands above the cut-off frequency.

20 4. The invention of claim 3, wherein the parametric audio encoding technique is BCC encoding.

5. The invention of claim 3, wherein the cut-off frequency is at least the effective audio bandwidth of the LFE channel.

6. The invention of claim 3, wherein the multi-channel audio signal is a 5.1 surround sound signal.

7. The invention of claim 1, further comprising transmitting the parametric audio codes for the first and second subsets of audio input channels.

25 8. An apparatus for encoding a multi-channel audio signal having a plurality of audio input channels, the apparatus comprising:

means for applying a parametric audio encoding technique to generate parametric audio codes for a first subset of the audio input channels for a first frequency region; and

means for applying the parametric audio encoding technique to generate parametric audio codes for a second subset of the audio input channels for a second frequency region, wherein:

- 5 the second frequency region is different from the first frequency region; and
 the second subset is different from the first subset.

9. A parametric audio encoder, comprising:

a downmixer adapted to generate one or more combined channels from a plurality of audio input channels of a multi-channel audio signal; and

10 an analyzer adapted to generate:

(1) parametric audio codes for a first subset of the audio output channels in a first frequency region; and

(2) parametric audio codes for a second subset of the audio output channels in a second frequency region, wherein:

- 15 the second frequency region is different from the first frequency region; and
 the second subset is different from the first subset.

10. The invention of claim 9, wherein the parametric audio codes are BCC codes.

11. The invention of claim 9, wherein:

20 the multi-channel audio signal is a surround sound signal having a plurality of regular channels and
at least one LFE channel;

the first subset includes all of the audio output channels;

the first frequency region corresponds to sub-bands at or below a specified cut-off frequency;

the second subset excludes the LFE channel; and

the second frequency region corresponds to sub-bands above the cut-off frequency.

25 12. The invention of claim 9, further the parametric audio encoder is adapted to transmit the
parametric audio codes for the first and second subsets of audio input channels.

13. A method for synthesizing a multi-channel audio signal having a plurality of audio output channels, the method comprising:

30 applying a parametric audio decoding technique to generate a first subset of the audio output
channels for a first frequency region; and

-11-

applying the parametric audio decoding technique to generate a second subset of the audio output channels for a second frequency region, wherein:

the second frequency region is different from the first frequency region; and
the second subset is different from the first subset.

5 14. The invention of claim 13, wherein the parametric audio decoding technique is BCC decoding.

15. The invention of claim 13, wherein:

the multi-channel audio signal is a surround sound signal having a plurality of regular channels and at least one LFE channel;

the first subset includes all of the audio output channels;

10 the first frequency region corresponds to sub-bands at or below a specified cut-off frequency;

the second subset excludes the LFE channel; and

the second frequency region corresponds to sub-bands above the cut-off frequency.

16. The invention of claim 15, wherein the parametric audio decoding technique is BCC decoding.

17. The invention of claim 15, wherein the cut-off frequency is at least the effective audio

15 bandwidth of the LFE channel.

18. The invention of claim 15, wherein the multi-channel audio signal is a 5.1 surround sound signal.

19. An apparatus for synthesizing a multi-channel audio signal having a plurality of audio output channels, the apparatus comprising:

20 means for applying a parametric audio decoding technique to generate a first subset of the audio output channels for a first frequency region; and

means for applying the parametric audio decoding technique to generate a second subset of the audio output channels for a second frequency region, wherein:

the second frequency region is different from the first frequency region; and

the second subset is different from the first subset.

25 20. A parametric audio decoder, comprising:

a parametric code processor adapted to generate parametric codes; and

a synthesizer adapted to apply the parametric codes to one or more combined channels to generate:

-12-

(1) a first subset of audio output channels of a multi-channel audio signal in a first frequency region; and

(2) a second subset of audio output channels of the multi-channel audio signal in a second frequency region, wherein:

5 the second frequency region is different from the first frequency region; and
 the second subset is different from the first subset.

21. The invention of claim 20, wherein the parametric codes are BCC codes.

22. The invention of claim 20, wherein:

10 the multi-channel audio signal is a surround sound signal having a plurality of regular channels and
 at least one LFE channel;
 the first subset includes all of the audio output channels;
 the first frequency region corresponds to sub-bands at or below a specified cut-off frequency;
 the second subset excludes the LFE channel; and
 the second frequency region corresponds to sub-bands above the cut-off frequency.