

METHOD AND APPARATUS FOR PHASE ESTIMATION IN A TRANSFORM CODER FOR HIGH QUALITY AUDIO

Technical Field

5

This invention relates in general to digital signal processing of audio signals, such as music signals. More particularly, the invention relates to the implementation of a high quality dual-channel digital audio encoder, based on the psychoacoustic model of the human auditory system, for digital storage or transmission.

10

Background Art

In order to more efficiently broadcast or record audio signals, the amount of information required to represent the audio signals may be reduced. In the case of digital audio signals, 15 the amount of digital information needed to accurately reproduce the original pulse code modulation (PCM) samples may be reduced by applying a digital compression algorithm, resulting in a digitally compressed representation of the original signal. The goal of the digital compression algorithm is to produce a digital representation of an audio signal which, when decoded and reproduced, sounds the same as the original signal, while using a minimum 20 of digital information for the compressed or encoded representation.

Recently, the use of psychoacoustic models in the design of audio coders has led to high compression ratios while keeping audible degradation in the compressed signal to a minimum. Description of one such method can be found in the Advanced Television Systems Committee 25 (ATSC) Standard document entitled "Digital Audio Compression (AC-3) Standard", Document A/52, 20 December, 1995. In the basic approach, the time domain signal is first converted to frequency domain using a bank of filters. Frequency domain masking of human auditory system is then exploited to maximise perceived fidelity of the signal transmitted at a given bit-rate.

30

- 2 -

Further compression can be successively obtained by use of a well known technique called coupling. Coupling takes advantage of the way the human ear determines directionality for very high frequency signals, in order to allow a reduction in the amount of data necessary to code an audio signal. At high audio frequency (approximately above 2KHz) the ear is
5 physically unable to detect individual cycles of an audio waveform, and instead responds to the envelope of the waveform. Consequently, the coder combines the high frequency coefficients of the individual channels to form a common coupling channel. The original channels combined to form the said coupling channel are referred to as coupled channels.

- 10 A basic encoder can form the coupling channel by simply taking the average of all the individual channel coefficients. A more sophisticated encoder can alter the sign of individual channels before adding them into the sum so as to avoid phase cancellations.

The generated coupling channel is next sectioned into a number of frequency bands.
15 Frequency sub-bands are grouped together to form coupling bands. For each such band and each coupled channel a coupling co-ordinate is transmitted to the decoder. To obtain the high frequency coefficients in any frequency band, for a particular coupled channel, from the said coupling channel, the decoder multiplies the coupling channel coefficients in that frequency band by the coupling co-ordinate of that channel for that particular frequency band. For a
20 dual channel implementation of such a decoder, a phase flag bit may also be provided for each coupled band of the coupling channel. A final step of phase-correction is then performed, by the decoder, in which the coefficients in each band are multiplied by the phase flag bit for that band.

- 25 The standard does not outline any specific method for determination of the phase flag bits. Ad hoc methods do exist but, due to their very nature, do not guarantee any assured performance, nor can be relied upon to provide minimum error between the original coefficients at encoder and the reconstructed, phase corrected, coefficients at the decoder.

- 3 -

Summary of the Invention

The phase flag for a band is, essentially, a function of the coefficients of the original channel and the coefficients of the coupling channel, in that band. Embodiments of the invention aim
5 to minimise the difference between the original coefficients at the encoder and the reconstructed coefficients at the decoder.

In accordance with the present invention, there is provided a method for computing a phase reconstruction coefficient in a dual channel digital audio encoder having first and second
10 encoded channels and a coupling channel, comprising computing transform coefficients for said first and second channels, computing coupling coefficients from the transform coefficients of the first and second channels, and computing a dot product of the corresponding transform coefficients for one of the first and second channels and the coupling coefficients, and determining the sign of the computed dot product.

15 The present invention also provides a method for computing a phase reconstruction coefficient in a dual channel digital audio encoder having first and second encoded channels and a coupling channel, comprising determining transform coefficients for one of the first and second channels, determining coupling coefficients from the first and second channels, and
20 computing the sign of the sum of corresponding transform and coupling coefficients over a predetermined frequency range of the coefficients.

Preferably the method includes computing a phase reconstruction coefficient for each of a plurality of coupling frequency bands for said one of the first and second channels.

25 Preferably the phase reconstruction coefficient computation is independent of a method used for computing the coupling coefficients, and independent of a method used for computing coupling coordinates for said first and second channels.

30 The present invention also provides a dual channel encoder for coding of audio information

- 4 -

which generates a coupling channel with at least one coupling band, comprising means for computing a dot product of input channel transform coefficients and coupling channel coefficients in said at least one coupling band, and means for determining the sign of said dot product for use as a phase flag corresponding to the at least one coupling band.

5

The present invention also provides a dual channel encoder for coding of audio information which generates a coupling channel and phase estimation data such that a difference between original coupled channel coefficients generated at the encoder and channel coefficients estimated at a compatible decoder have a least square error.

10

There is further provided, in accordance with the present invention, an encoder for transform coding digital audio information from first and second channels, the encoder producing a coupling channel arranged in at least one frequency band and at least one phase flag corresponding to the at least one frequency band, wherein the at least one phase flag is

15 computed according to:

$$\text{phase flag} = \text{sign} (\sum (b_i * c_i))$$

wherein

b_i represents transform coefficients for one of the first and second channels,

c_i represents transform coefficients for the coupling channel, and

20 index i extends over the frequency range of the band.

Brief Description of the Drawings

The invention is described in greater detail hereinbelow, by way of example only, through
25 description of an embodiment thereof and with reference to the accompanying drawings,
wherein:

Figure 1 is a diagrammatic illustration of an audio encoding/decoding process including channel coupling and phase reconstruction;

Figures 2a and 2b are vector space diagrams; and

30 Figure 3 is a block diagram of a dual channel audio encoder apparatus constructed

- 5 -

according to an embodiment of the invention.

Detailed Description of the Preferred Embodiments

5 As mentioned, one method for compression of digital audio signals is described in the ATSC Standard entitled "Digital Audio Compression (AC-3) Standard", (Document A/52, 20 December, 1995), and the contents of that document are expressly incorporated herein by reference. The following description of a preferred embodiment of the invention is presented in the context of an audio encoder which is compatible with the AC-3 Standard.

10

An AC-3 encoder accepts PCM audio and produces an encoded bit stream consistent with the AC-3 standard. The specifics of the audio encoding process are not normative requirements of the standard, but a compatible encoder must produce a bit stream matching the syntax set out in the AC-3 standard document, which, when decoded as specified, produces audio of
15 sufficient quality for the intended application.

The AC-3 algorithm achieves high coding gain (the ratio of the input bit-rate to the output bit-rate) by coarsely quantizing a frequency domain representation of the audio signal. The basic encoding process is as follows. The first step in the encoding process is to transform the
20 representation of audio from a sequence of PCM time samples into a sequence of blocks of frequency coefficients. This is done in the analysis filter bank. Overlapping blocks of 512 time samples are multiplied by a time window and transformed into the frequency domain. Due to the overlapping blocks, each PCM input sample is represented in two sequential transformed blocks. The frequency domain representation may then be decimated by a factor
25 of two so that each block contains 256 frequency coefficients. The individual frequency coefficients are represented in binary exponential notation as a binary exponent and a mantissa. The set of exponents is encoded into a coarse representation of the signal spectrum which is referred to as the spectral envelope. This spectral envelope is used by a core bit allocation routine which determines how many bits to use to encode each individual mantissa.
30 The spectral envelope and the coarsely quantized mantissas for 6 audio blocks (1536 audio

- 6 -

samples) are formatted into an AC-3 frame. The AC-3 bit stream is a sequence of AC-3 frames.

Channel coupling is performed during encoding by averaging certain transform coefficients
5 across channels that are included in the coupling channel. Each coupled channel has a unique set of coupling coordinates which are used to preserve the high frequency envelopes of the original channels. The coupling process is performed for transform coefficients corresponding to frequencies above a coupling frequency that is defined by a pre-specified value, *cplbegf*.

10

The decoder converts the coupling channel back into individual channels by multiplying the coupled channel transform coefficient values by the coupling coordinate for that channel and frequency band. An additional processing step occurs where only two channels are coupled and phase flag bits are employed. In this case a phase-flag-in-use bit (*phsflginu*) is set in the
15 encoded bit-stream, and phase restoration bits are retrieved from the bit-stream by the decoder in the form of phase flag bits. The phase flag bits are arranged to represent the coupling bands in a frequency ascending order. If a phase flag bit is set for a particular band, all of the right channel transform coefficients within the coupled band are negated (inverted) after modification by the coupling coordinate, but before inverse transformation.

20

Transform coefficients #37 through #252 are grouped into 18 sub-bands of 12 coefficients each, as shown below in Table 1. The parameter *cplbegf* indicates the number of the coupling sub-band which is the first to be included in the coupling process. Below the frequency (or transform coefficient number, *tc#*) indicated by *cplbegf* all channels are
25 independently coded. Above the frequency indicated by *cplbegf*, channels included in the coupling process share the common coupling channel, up to the frequency (or *tc#*) indicated by a parameter *cplendf*. The coupling channel is coded up to the frequency (or *tc#*) indicated by *cplendf*, which indicates the last coupling sub-band which is coded. The parameter *cplendf* is interpreted by adding 2 to its value, so the last coupling sub-band which is coded
30 can range from 2-17.

- 7 -

	coupling sub-band#	low tc#	high tc#	If cutoff (kHz) @fs=48 kHz	If cutoff (kHz) @fs=48 kHz	If cutoff (kHz) @fs=48 kHz	If cutoff (kHz) @fs=48 kHz
	0	37	48	3.42	4.55	3.14	4.18
5	1	49	60	4.55	5.67	4.18	5.21
	2	61	72	5.67	6.80	5.21	6.24
	3	73	84	6.80	7.92	6.24	7.28
	4	85	96	7.92	9.05	7.28	8.31
	5	97	108	9.05	10.17	8.31	9.35
10	6	109	120	10.17	11.30	9.35	10.38
	7	121	132	11.30	12.42	10.38	11.41
	8	133	144	12.42	13.55	11.41	12.45
	9	145	156	13.55	14.67	12.45	13.48
	10	157	168	14.67	15.80	13.48	14.51
15	11	169	180	15.80	16.92	14.51	15.55
	12	181	192	16.92	18.05	15.55	16.58
	13	193	204	18.05	19.17	16.58	17.61
	14	205	216	19.17	20.30	17.61	18.65
	15	217	228	20.30	21.42	18.65	19.68
20	16	229	240	21.42	22.55	19.68	20.71
	17	241	252	22.55	23.67	20.71	21.75

The coupling channel sub-bands are combined in to coupling bands for which coupling coordinates are generated for each coupled channel (and included in the bit stream). The coupling band structure is indicated by a parameter *cplbndstrc[sbnd]*. Each bit of the *cplbndstrc[]* array indicates whether the sub-band corresponding to the index is combined into the previous (lower in frequency) coupling band. Coupling bands are thus each structured based on an integral number of coupling channel sub-bands.

Figure 1 diagrammatically illustrates the general coupling process. In this Figure, the Figure portions 1a to 1i illustrate diagrammatically the various frequency coefficients, coupling

- 8 -

coordinates and phase flag values as functions of frequency over a portion of an encoding followed by decoding process. In a dual channel encoder, the input bit stream comprises interleaved digitised signals of two channels, the two individual channels hereinafter being identified as a (left channel) and b (right channel). The first step is to extract a block of the
 5 digitised samples, for both channel a and b , from the input stream and transform it to the frequency domain using filter banks, to generate frequency domain coefficients. The general frequency domain coefficients for channels a and b are hereinafter identified as sequences a_i and b_i (see Fig. 1a and 1b), respectively, where the subscript i identifies the corresponding frequency sequence number of the coefficient.

10

A frequency point called $f_{cplbegf}$ (coupling begin frequency) is selected by the encoder based on the $cplbegf$ parameter as outlined above. Frequency coefficients a_i and b_i , for all i below $f_{cplbegf}$ are sent individually by the encoder into the compressed bit stream (see Fig. 1c and 1h).

15 Frequency coefficients a_i and b_i , for all i above $f_{cplbegf}$ are combined by a binary operator to generate a new channel called the coupling channel, c . Channels a and b which are used to form the coupling channel, are referred to as coupled channels. Any binary operator ($\oplus$) can be used for generation of the coupling channel, its exact nature not being specified by the AC-3 standard and independent of the phase determination strategy described herein. Thus, the
 20 method and apparatus described herein according to embodiments of the present invention, for estimation of phase reconstruction data, place no restriction on the coupling channel generation strategy.

The coupling channel, c , comprises frequency domain coefficients which are identified as c_i ,
 25 $= a_i \oplus b_i$. The coupling channel is sectioned into a number of consecutive bands (see Fig. 1e), the encoder controlling the coupling band structure (the bandwidth of each coupled band) as outlined above and described in the AC-3 standard. For each such band, and each coupled channel, a coupling co-ordinate value is generated by the encoder (see Fig. 1d and 1f).

30 To restore the original channel coefficients a_i or b_i in a particular band, the decoder must

- 9 -

multiply each coupling channel coefficient c_i in that band by the coupling-coordinate value for that band and channel (see Fig. 1d and 1e). For example, if ψ is the coupling co-ordinate for a band in the coupled channel b , the coefficients estimate for b at the decoder, called $\hat{b}_i$, will be

5

$$\hat{b}_i = \psi c_i \quad \text{Eq. 1}$$

The encoder measures the original signal power in the band for the individual coupled channel, as well as the power in the coupling channel for the same band. The ratio of
10 original signal power within a band to the coupling channel power in the same band is termed as the coupling co-ordinate for the band. However, once again, the phase estimation method described herein is independent of the coupling co-ordinate generation strategy.

In a dual channel encoder, for the second (right) channel i.e. channel b , a phase flag
15 information is transmitted for each coupling band to allow phase correction at the decoder (see Fig. 1g). Therefore, Eq. 1 can be rewritten as

$$\hat{b}_i = \text{phase_flag} * \psi c_i \quad \text{Eq. 2}$$

20 A difficulty is how to determine the phase flags. A simple approach could be to compute the sum of the channel b -coefficients in the band, i.e. $\sum b_i$, and also the sum of the coefficients in the band for the coupling channel, i.e. $\sum c_i$. If both the sums are of the same sign, then phase flag is set to +1, else if they are of opposite signs, then phase flag is set to -1. That is,

25

$$\text{phase_flag} = \text{sign}[(\sum b_i) * (\sum c_i)] \quad \text{Eq. 3}$$

However, this approach can lead to spurious results, and to gain insight into the functioning of the phase flag, a more detailed analysis is required.

30

- 10 -

Channel coefficients for a particular band can be visualised as vectors, in a n -dimensional space, n being the number of frequency coefficient points in the band. Thus, $\vec{b} = (b_1, b_2, \dots, b_n)$ and $\vec{c} = (c_1, c_2, \dots, c_n)$. Figures 2a and 2b illustrate the vector space formed by the channel coefficients in a band. The estimate of vector $\vec{b}$ at decoder can be,

5

$$\vec{b}_1 = \psi \vec{c} \quad \text{Eq. 4}$$

or

$$\vec{b}_2 = -\psi \vec{c} \quad \text{Eq. 5}$$

depending whether phase flag is sent as +1 or -1, respectively.

10

In Fig. 2a, xx' is defined as the hyper-plane perpendicular to vector $\vec{c}$, dividing the n -dimensional vector space into two parts. When $\vec{b}$ lies on the same side of the hyper-plane as $\vec{c}$, then $\vec{b}_1$ is a better estimate of $\vec{b}$ than $\vec{b}_2$ (as in the case shown in Fig. 2a), i.e.

$||\vec{b} - \vec{b}_1|| < ||\vec{b} - \vec{b}_2||$. Similarly, when $\vec{b}$ lies on opposite side of the hyper-

15

plane as $\vec{c}$, then $\vec{b}_2$ is a better approximation than $\vec{b}_1$.

Adopting this approach, it is then apparent from Figures 2a and 2b that if the dot product ($\cdot$) or the *cosine* of the angle between the vectors $\vec{b}$ and $\vec{c}$ is positive, i.e. the angle between them is less than 90° , then phase flag should be set as +1, otherwise it should be -1. A
20 statement of the method can be expressed as follows:

$$\text{phase_flag} = \text{sign}(\vec{b} \cdot \vec{c}) = \text{sign}(\sum(b_i * c_i)) \quad \text{Eq. 6}$$

Having noted the geometric interpretation of the phase flag determination method, it is then
25 possible to analyse it from the perspective of least-square-error between the original channel coefficient b_i at encoder and the estimated channel coefficients $\hat{b}_i$ at decoder.

Assuming, $\text{sign}[\sum(b_i * c_i)] = +1$, it is then necessary to show that ψc_i is a closer approximate than $-\psi c_i$. That is,

30

- 11 -

$$\sum (b_i - \psi c_i)^2 < \sum (b_i + \psi c_i)^2$$

$$\Rightarrow [\sum b_i^2 - 2\psi \sum b_i c_i + \psi^2 \sum c_i^2] < [\sum b_i^2 + 2\psi \sum b_i c_i + \psi^2 \sum c_i^2]$$

Cancelling identical terms on both sides of the inequality,

$$\Rightarrow -2\psi \sum b_i c_i < 2\psi \sum b_i c_i$$

$$5 \quad \Rightarrow \sum b_i c_i > 0$$

which is true under the assumption.

It can therefore be observed that Eq. 6, gives phase flag with guaranteed least-square-error between the original channel coefficients at encoder and the estimated channel coefficients at
10 decoder.

In summary, assuming that the frequency coefficients within a particular coupling band are identified as:

a_i , for the first coupled channel,

15 b_i , for the second coupled channel,

c_i , for the coupling channel,

and index i extends over the frequency range of the band. Then, the phase flag data for the coupling band can be computed according to:

$$phase_flag = sign(\sum (b_i * c_i)) \quad \text{Eq. 7}$$

20

Note: the $sign()$ function has the general interpretation, that is, it returns a +1 if the function parameter is a positive number and -1 if it is negative. For the case of zero, any value +1 or -1, can be considered as the signals are then essentially orthogonal (uncorrelated).

25 Figure 3 is a block diagram showing a dual channel audio encoder apparatus 10 for implementing this phase estimation strategy. The encoder 10 receives PCM coded audio samples corresponding to left and right audio channels referred to as channel a (12) and channel b (14), respectively. The audio data from channels a and b are passed to respective analysis filter banks 16, 18 of conventional form, which transform the audio data into
30 frequency domain transform coefficients a_i and b_i . The transform coefficients are passed to

- 12 -

a coupling processor 20 which generates coupling channel coefficients c_i .

The encoder 10 includes a coupling band determination processor 22, which determines how the bands in the coupling channel are structured according to predefined parameters (i.e. determines a correspondence between the coupling channel coefficients and the coupled channels' coupling coordinates). The banding structure processor 22 controls left and right channel coupling coordinate computation processors 24, 26. The coupling coordinate computation processors receive as input the coupling channel coefficients c_i and the respective left or right channel transform coefficients a_i , b_i , and generate respective channel coupling coordinates according to the coupling band structure dictated by the banding structure processor 22.

A phase estimation processor 28 is also provided which receives as input the coupling channel coefficients c_i and the right channel coefficients b_i . The phase estimation processor 28 is also controlled by the banding structure processor 22 so that the phase flags generated thereby correspond to the predefined frequency bands of the channel coupling coordinates. For each coupling channel frequency band the phase estimation processor computes a corresponding phase flag according to the operation outlined by Eq. 7, above. The phase flags are output to be included in the encoded bit stream along with the coupling channel coefficients c_i , and the left and right channel coupling coordinates.

The dual channel audio encoder apparatus 10 is preferably implemented as part of an application specific integrated circuit, in which case the functional blocks described above may comprise circuit portions which interact with one another under control of timing and control circuitry. Alternatively, however, the encoder apparatus may be implemented by a more general purpose data processing circuit such as a conventional microprocessor, and in that case the functions performed by the individual functional blocks as above described may all be performed by the same microprocessor circuit according to instructions defining the functions which are stored as, for example, micro-code.

- 13 -

The foregoing detailed description of embodiments of the invention have been presented by way of example only, and this description is not intended to be considered limiting to the invention as defined in the claims appended hereto.



- 14 -

Claims:

1. A method for computing a phase reconstruction coefficient in a dual channel digital audio encoder having first and second encoded channels and a coupling channel, comprising
5 computing transform coefficients for said first and second channels, computing coupling coefficients from the transform coefficients of the first and second channels, and computing a dot product of the corresponding transform coefficients for one of the first and second channels and the coupling coefficients, and determining the sign of the computed dot product.
- 10 2. A method as claimed in claim 1, including computing a phase reconstruction coefficient for each of a plurality of coupling frequency bands for said one of the first and second channels.
3. A method as claimed in claim 1 or 2 which is independent of a method used for
15 computing the coupling coefficients.
4. A method for computing a phase reconstruction coefficient in a dual channel digital audio encoder having first and second encoded channels and a coupling channel, comprising determining transform coefficients for one of the first and second channels, determining
20 coupling coefficients from the first and second channels, and computing the sign of the sum of corresponding transform and coupling coefficients over a predetermined frequency range of the coefficients.
5. A method as claimed in claim 4, including computing a plurality of phase
25 reconstruction coefficients for different predetermined frequency bands.
6. A method as claimed in claim 4 or 5 which is independent of a method used for computing the coupling coefficients.
- 30 7. A method as claimed in claim 6 which is independent of a method used for computing



coupling coordinates for said first and second channels.

8. A dual channel encoder for coding of audio information which generates a coupling channel with at least one coupling band, comprising means for computing a dot product of
5 input channel transform coefficients and coupling channel coefficients in said at least one coupling band, and means for determining the sign of said dot product for use as a phase flag corresponding to the at least one coupling band.
9. A dual channel encoder for coding of audio information which generates a coupling
10 channel and phase estimation data such that a difference between original coupled channel coefficients generated at the encoder and channel coefficients estimated at a compatible decoder have a least square error.
10. An encoder for transform coding digital audio information from first and second
15 channels, the encoder producing a coupling channel arranged in at least one frequency band and at least one phase flag corresponding to the at least one frequency band, wherein the at least one phase flag is computed according to:
- $$\text{phase flag} = \text{sign} (\sum (b_i * c_i))$$
- wherein
- 20 b_i represents transform coefficients for one of the first and second channels,
 c_i represents transform coefficients for the coupling channel, and
index i extends over the frequency range of the band.