

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5186543号
(P5186543)

(45) 発行日 平成25年4月17日(2013.4.17)

(24) 登録日 平成25年1月25日(2013.1.25)

(51) Int.Cl. F I
G 1 0 L 19/00 (2013.01) G 1 0 L 19/00 2 1 3
H 0 4 S 5/02 (2006.01) H 0 4 S 5/02 Y

請求項の数 3 (全 14 頁)

(21) 出願番号	特願2010-236053 (P2010-236053)	(73) 特許権者	506427990
(22) 出願日	平成22年10月21日(2010.10.21)		ドルビー・インターナショナル・アクチボラゲット
(62) 分割の表示	特願2005-289552 (P2005-289552) の分割		DOLBY INTERNATIONAL AB
原出願日	平成14年7月10日(2002.7.10)		オランダ、1101 セー・エン アムステルダム・ザイドースト、ヘリケルベルク
(65) 公開番号	特開2011-34102 (P2011-34102A)		ベーク、1-35、アポロ・ビルディング、3・エー
(43) 公開日	平成23年2月17日(2011.2.17)	(74) 代理人	100112715
審査請求日	平成22年10月21日(2010.10.21)		弁理士 松山 隆夫
(31) 優先権主張番号	0102481-9	(72) 発明者	フレドリック・ヘン
(32) 優先日	平成13年7月10日(2001.7.10)		スウェーデン国 S-16831 ブロンマ リタルヴェーゲン 14
(33) 優先権主張国	スウェーデン(SE)		
(31) 優先権主張番号	0200796-1		
(32) 優先日	平成14年3月15日(2002.3.15)		
(33) 優先権主張国	スウェーデン(SE)		

最終頁に続く

(54) 【発明の名称】 低ビットレートオーディオ符号化用の効率的かつスケーラブルなパラメトリックステレオ符号化

(57) 【特許請求の範囲】

【請求項1】

一次ビットストリーム部分と二次ビットストリーム部分を含む符号化されたビットストリームを復号化する復号器であって、

低帯域コア復号器信号と、2つのチャンネルをもつ信号の周波数帯域における全パワーを表すレベルパラメータとを得るために、前記符号化されたビットストリームの前記一次ビットストリーム部分を逆多重化し、2つのチャンネルをもつ信号の各周波数帯域に対するバランスパラメータを得るために、前記符号化されたビットストリームの前記二次ビットストリーム部分を逆多重化するデマルチプレクサと、

低帯域モノラル信号又は低帯域ステレオ信号をもつ低帯域出力信号を生成する低帯域コア復号器と、

前記低帯域出力信号を使用して合成高帯域を生成し、前記レベルパラメータおよび前記バランスパラメータによって前記合成高帯域を調整して2つのチャンネルの調整された合成高帯域を取得し、かつその調整された合成高帯域と前記低帯域出力信号を結合して広帯域出力信号を得る高周波再構成装置とを備え、

前記広帯域出力信号は、前記広帯域出力信号の低帯域部分における低帯域出力信号と、前記広帯域出力信号の高帯域部分における調整された合成高帯域とを備える復号器。

【請求項2】

前記ビットストリームは前記2つのチャンネルの和信号を含み、前記低帯域コア復号器はモノラル信号としての低帯域出力信号を生成するように動作するものである請求項1に記

10

20

載の復号器。

【請求項 3】

一次ビットストリーム部分と二次ビットストリーム部分を含む符号化されたビットストリームを復号化する方法であって、

低帯域コア復号器信号と、2チャンネルをもつ信号の周波数帯域における全パワーを表すレベルパラメータとを得るために、前記符号化されたビットストリームの前記一次ビットストリーム部分を逆多重化するステップと、

2つのチャンネルをもつ信号の各周波数帯域に対するバランスパラメータを得るために、前記符号化されたビットストリームの前記二次ビットストリーム部分を逆多重化するステップと、

低帯域モノラル信号又は低帯域ステレオ信号をもつ低帯域出力信号を生成するために前記低帯域コア復号器信号を復号化するステップと、

前記低帯域出力信号を用いて高周波再構成により合成高帯域を生成し、前記レベルパラメータおよび前記バランスパラメータによって前記合成高帯域を調整して2つのチャンネルの調整された合成高帯域を取得し、かつ前記調整された合成高帯域及び低帯域出力信号を結合して広帯域出力信号を得るステップとを備え、

前記広帯域出力信号は、前記広帯域出力信号の低帯域部分における前記低帯域出力信号と、前記広帯域出力信号の前記高帯域部分における前記調整された合成高帯域とを備える方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、低ビットレートオーディオ情報源符号化システムに関する。入力信号のステレオ特性の様々なパラメトリック表現を紹介し、スペクトル包絡線の擬似ステレオからフルステレオ符号化にわたって、復号器側におけるその適用について説明する。後者はHFR（高周波再構成）に基づくコーデックに特に適している。

【背景技術】

【0002】

オーディオ情報源符号化技術は2種類、すなわち自然オーディオ符号化（natural audio coding）及び会話符号化（speech coding）に分類することができる。中ないし高ビットレートでは、一般に自然オーディオ符号化が会話信号にも音楽信号に使用され、ステレオ伝送とステレオ再生が可能である。低ビットレートしか利用できない適用例、例えば低速電話モデム接続のユーザを対象とするインターネットストリーミングオーディオ、または新興のデジタルAM放送システムでは、オーディオ番組素材のモノラル符号化は避けられない。しかし、それでもなおステレオの感覚が望ましい。ヘッドホンで聞くとときに純粋なモノラル信号が「頭の中」から生じているように知覚され、それが不快な経験であり得る場合は、特にそうである。

【0003】

この問題に取り組む1つの手法は、復号器の側で、受信した純粋なモノラル信号からステレオ信号を合成することである。長年にわたり、幾つかの異なる「疑似ステレオ」生成器が提案されてきた。例えば[米国特許第5,883,962号]には、信号の遅延/移相バージョンを未処理信号に加えることによってモノラル信号を強化し、それによって擬似的なステレオを生み出すことが記載されている。そこでは、2つのチャンネルが後で信号経路で加えられれば強化信号が確実に相殺されるようにしながら、処理された信号を同レベルであるが逆の符号をもって2つの出力の各々の原信号に加算する。[PCT WO 98/57436]には、強化信号の上記のモノラル両立性（mono-compatibility）はないが、同様のシステムが示されている。先行技術の方法は、それらが純粋な後処理として適用されるという点で共通している。言い換えると、ステレオサウンドステージの位置はあるか、ステレオ幅の程度に関する情報さえも復号器には得られない。したがって、疑似ステレオ信号が原信号のステレオ特性との類似性をもつかどうかは分からない。先行技術の

10

20

30

40

50

システムが不十分であることが顕著に現れるのは、原信号が純粋なモノラル信号である場合であり、それは往々にして会話録音の場合に当てはまる。このモノラル信号は復号器で盲目的に合成ステレオ信号に変換され、それは会話の場合、しばしば耳障りなアーチファクトを引き起こし、明瞭さと会話の了解度を低減することがある。

【0004】

低ビットレートの真のステレオ伝送を目的とする他の先行技術のシステムは、一般的に和差符号化方式を採用している。例えば、元の左側信号(L)と右側信号(R)は和信号 $S = (L + R) / 2$ と差信号 $D = (L - R) / 2$ に変換され、その後符号化されて伝送される。受信機は、演算 $L = S + D$ と $R = S - D$ により元のL/R信号が再生されるように、S信号とD信号を復号化する。この利点は、LとRとの間の冗長度を使用できることが非常に多く、それにより符号化されるDに含まれる情報がSに含まれる情報より少なくなつて必要なビット数が少なくてすむことである。極端な例は純粋なモノラル信号の場合、すなわちLとRが同一の場合であることは明らかである。従来のL/Rコーデックはこのモノラル信号を2回符号化するのに対し、S/Dコーデックはこの冗長度を検出し、D信号は(理想的には)ビットを全く必要としない。別の極端な例は、「位相ずれ」(out of phase)信号に相当する $R = -L$ という状況で現れる。ここではD信号はLと計算されるのに対し、S信号は零である。この場合もS/D方式は標準的L/R符号化より明らかな利点を有する。しかし、送信中に例えば $R = 0$ となる状況を考える。それはステレオ録音が始まった初期の頃には珍しくはなかった。その状況下では、SとDはどちらもL/2に等しく、S/D方式は何らの利点ももたらさない。それに対して、L/R符号化はこれを非常にうまく処理する。というのは、R信号はビットを必要としないからである。このため、先行技術のコーデックは、所定の瞬間にどちらの方法を使用するのが最も有利であるかによって、これらの2つの符号化方式の間の適応型切換えを採用している。上の極端な例は、(会話のみの番組に使用されるデュアルモノラルを除いて)単なる理論である。したがって、実世界のステレオ番組素材はかなりの量のステレオ情報を含み、たとえ上記切換えが実現されても、結果的に得られるビットレートは多くの用途には往々にして高すぎる。さらに、上記の再合成関係から分かる通り、量子化誤差はL信号とR信号における無視できないレベルの誤差に変わるので、ビットレートをさらに低減しようとしてD信号の非常に粗い量子化を実行することは不可能である。

【発明の開示】

【0005】

本発明は、符号化と伝送の前に、信号のステレオ特性の検出を採用する。最も単純な形では、入力ステレオ信号に存在するステレオ配合(stereo perspective)の量を検出器で測定する。次いでこの量をステレオ幅パラメータとして、原信号の符号化されたモノラル和と共に伝送する。受信機はモノラル信号を復号化し、疑似ステレオ生成器を使用して、前記パラメータによって制御されるステレオ幅の適切な量を適用する。特別な場合として、モノラル入力信号は零ステレオ幅として符号化され、したがって復号器でステレオ合成は適用されない。本発明では、ステレオ幅の有用な尺度を、例えば元の左右チャンネルの差信号または相互相関から導出することができる。そのような計算の値を少数の状態で表現することができ、それらは適当な時間間隔で、あるいは必要に応じて伝送される。本発明はまた、一般的に低ビットレート符号化信号と関連のある脱マスクング(unmasking)符号化アーチファクトの危険性を低減するために、合成ステレオ成分をフィルタリングする方法をも教示する。

【0006】

また、ステレオフィールドにおける全体的ステレオバランスまたは位置を符号器で検出する。この情報をバランスパラメータとして、符号化されたモノラル信号と共に、必要に応じて上記の幅パラメータと一緒に、効率的に伝送する。したがって、サウンドステージのいずれかの側への変位は、2つの出力チャンネルの利得を対応づけて変更することによって、復号器で再生することができる。本発明では、このステレオバランスパラメータは左側信号パワーと右側信号パワーとの商から導き出すことができる。両パラメータの伝送は

10

20

30

40

50

フルステレオ符号化と比較して必要なビット数が非常に少なく済み、それにより総ビットレート要求が低く維持される。より正確なパラメトリックステレオ描写を提供する本発明のより精巧な形態では、それぞれが別個の周波数帯を表わす幾つかのバランスパラメータとステレオ幅パラメータが使用される。

【 0 0 0 7 】

周波数帯域毎の作用に一般化されるバランスパラメータは、左側信号パワーと右側信号パワーの和として計算されるレベルパラメータの対応する帯域毎の作用とともに、ステレオ信号のパワースペクトル密度についての新しく、詳細の程度が任意な表現を可能にする。S/Dシステムも備えているステレオ冗長度からの利点以外に、この表現の特に有利な点は、ステレオスペクトル包絡線に戻したときに量子化誤差がレベルの誤差ではなくむしろ「空間誤差」すなわちステレオパノラマで知覚される位置の誤差になるので、バランス信号を同様のレベルより低い精度で量子化できることである。総合信号がいずれか一方のチャンネルにひどく偏位しているときは、レベル/バランス方式も従来の切換え式のL/R及びS/Dシステムと同様に、より効率的なレベルL/レベルR信号に適応的に切り換えることができる。上記のスペクトル包絡線符号化方式は、パワースペクトル包絡線の効率的な符号化が要求されるときはいつでも使用することができ、新しいステレオソースコーデックにツールとして組み込むことができる。特に興味深い適用は、原信号の高帯域包絡線に関する情報によって誘導されるHFRシステムである。そのようなシステムでは、低帯域は任意のコーデックによって符号化と復号化が行なわれ、高帯域は復号化された低帯域信号と伝送された高帯域包絡線情報を使用して復号器で再生される [P C T W O 9 8 / 5 7 4 3 6]。さらに、包絡線符号化をレベル/バランス動作に固定することによって、スケラブルなHFRに基づくステレオコーデックを形成することを可能にする。これにより、レベル値は一次ビットストリーム内に供給され、それは装置によって異なるが一般的にモノラル信号に復号化される。バランス値は、一例としてI B O C (In-Band On-Channel) デジタル A M 放送システムをとりあげると、一次ビットストリームに加えて、送信機に近接する受信機に利用可能な二次ビットストリーム内に供給される。これらの2つのビットストリームが結合されると、復号器はステレオ出力信号を生成する。一次ビットストリームは、レベル値に加えてステレオパラメータ、例えば幅パラメータを含むことができる。したがって、このビットストリームのみでの復号化がすでにステレオ出力を生じており、それは両方のビットストリームが得られるときに改善される。

【 図面の簡単な説明 】

【 0 0 0 8 】

【 図 1 】パラメトリックステレオ符号器モジュールによって強化された符合器及びパラメトリックステレオ復号器モジュールによって強化された復号器を含む情報源符号化システムを示す図である。

【 図 2 a 】パラメトリックステレオ復号器モジュールのブロック略図である。

【 図 2 b 】制御パラメータ入力を持つ疑似ステレオ生成器のブロック略図である。

【 図 2 c 】制御パラメータ入力を持つバランス調整器のブロック略図である。

【 図 3 】マルチバンドバランス調整と組み合わせられたマルチバンド疑似ステレオ生成を使用するパラメトリックステレオ復号器モジュールのブロック略図である。

【 図 4 a 】スペクトル包絡線のレベル/バランス符号化を使用する、スケラブルなHFRに基づくステレオコーデックの符号器側のブロック略図である。

【 図 4 b 】対応する復号器側のブロック略図である。

【 発明を実施するための最良の形態 】

【 0 0 0 9 】

次に、本発明を添付の図面を参照しながら、本発明の範囲または精神を限定しない実施例によって説明する。

【 0 0 1 0 】

以下で述べる実施形態は、本発明の原理の単なる説明である。ここに記載する構成及び詳細説明の変化例ならびに変形例が、当業者には明らかであることを理解されたい。した

10

20

30

40

50

がって、本書における実施形態の記述及び説明によって提示される特定の詳細によってではなく、特許請求の範囲によってのみ限定するつもりである。分かりやすくするために、以下の例は全て2つのチャンネルシステムを想定しているが、当業者には明白であるように、当該方法は5.1システムなどのマルチチャンネルシステムにも適用できる。

【0011】

図1は、符号器107及び復号器115を備えた任意の情報源符号化システムを、符号器及び復号器がモノラルモードで動作する場合に、本発明によるパラメトリックステレオ符号化によっていかに強化することができるかを示す。L及びRは、左側アナログ入力信号及び右側アナログ入力信号を表わすことにする。これらはAD変換器101に供給される。AD変換器からの出力はダウンミックス105でモノラル信号に変換され、そのモノラル信号は符号器107で符号化される。また、ステレオ信号は、後述する1つまたは幾つかのステレオパラメータを計算するパラメトリックステレオ符号器103に送られる。これらのパラメータはマルチプレクサ109によって符号化モノラル信号と結合され、ビットストリーム111を形成する。ビットストリームは格納または伝送され、その後復号器側でデマルチプレクサ113によって抽出される。モノラル信号は復号器115で復号化され、ステレオパラメータ117を制御信号として使用するパラメトリックステレオ復号器119によってステレオ信号に変換される。最後に、ステレオ信号はアナログ出力L'及びR'を供給するDA変換器121に送られる。図1のトポロジーは、以下においてより簡単なバージョンから始めて、詳述する1組のパラメトリックステレオ符号化方法に共通である。

【0012】

本発明に係るステレオ特性のパラメータ化の1つの方法は、符号器側で原信号のステレオ幅を決定することである。大まかに言うと、LとRとの間の類似度が高ければ算出されるDの値は小さくなり、その逆もしかりであるので、ステレオ幅の最初の概算は差信号 $D = L - R$ である。特殊な例はデュアルモノラルであり、この場合 $L = R$ であり、したがって $D = 0$ である。したがって、この簡単なアルゴリズムでも、疑似ステレオが望ましくないニュース放送に一般に関連付けられる、モノラル入力信号の型を検出することができる。しかし、異なるレベルでL及びRに供給されるモノラル信号は、知覚される幅が零であっても、零D信号を生じない。したがって、実際には、例えば相互相関法を使用する、より精巧な検出器が必要になるであろう。レベルに依存しない検出器を達成するためには、何らかの方法で左右の差または相関を記述する値を全信号レベルにより正規化しなければならない。上述の検出器についての問題は、会話から音楽へ/音楽から会話への遷移中に、モノラル会話がずっと弱いステレオ信号、例えばステレオ雑音又は背景音楽と混合される場合である。会話が中断すると、検出器は次いでワイドステレオ信号を示す。これは、ステレオ幅値を前の全エネルギーレベルの情報を含む信号、例えば全エネルギーのピーク減衰信号で正規化することによって解決される。さらに、ステレオ幅検出器が高周波雑音又はチャンネルの異なる高周波ひずみによってトリガされることを防止するために、検出器信号は、一般的に音声の第2フォルマントより上寄りのカットオフ周波数の低域フィルタによって事前フィルタリングする必要があり、また必要に応じて、アンバランス信号オフセットまたはハムを回避するために高域フィルタによっても、事前フィルタリングする必要がある。検出器の型に関係なく、算出されたステレオ幅は、モノラルからワイドステレオまでの範囲全体を網羅する有限集合の値で表現される。

【0013】

図2aは、図1に導入したパラメトリックステレオ復号器の中身の一例を示す。パラメータBによって制御される「バランス」と表示されたブロック211については後述するが、とりあえずは迂回するものとする。「幅」と表示されたブロック205はモノラル入力信号を受け、ステレオ幅の感覚を再現する。ここでは幅の量はパラメータWによって制御される。任意のパラメータS及びDについては後述する。本発明では、低周波数範囲を「厳密」に、かつ影響されない状態に維持するために、低域フィルタ203及び高域フィルタ201からなるクロスオーバーフィルタを組み込むことによって、主観的により優れ

10

20

30

40

50

た音質を達成することができる。これにより、高域フィルタからの出力のみが幅ブロックに送られる。幅ブロックからのステレオ出力は、加算器 207 及び 209 によって低域フィルタからのモノラル出力に加算され、ステレオ出力信号を形成する。

【0014】

幅ブロックには、背景技術のところで示したものの、シュレーダ(Schroeder)型の初期の反射シミュレーション装置(マルチタップ遅延)またはリバーブレータ(reverberator)など、任意の先行技術疑似ステレオ生成器を使用することができる。図2bは、モノラル信号Mが供給される疑似ステレオ生成器の一例を示す。ステレオ幅の量は増幅器215の利得によって決定され、この利得はステレオ幅パラメータWの関数である。利得が高ければ高いほど、ステレオ感覚が広くなり、零利得は純粋なモノラル表現に対応する。増幅器215からの出力は遅延(221)され、正負の符号を使用して、2つの直接信号に加算(223及び225)される。ステレオ幅を変えたときに全体的な再生レベルを著しく変化させないために、直接信号の補償減衰を組み込む(213)ことができる。例えば、遅延された信号の利得がGである場合、直接信号の利得は $(1 - G^2)$ の平方根として選択することができる。本発明では、高周波ロールオフ(roll-off)を遅延信号経路217に組み込むことができ、これは疑似ステレオから生じる符号化アーチファクトの脱マスクングを回避する。図2a及び図2bに信号X、S及びDとして示されるように、必要に応じてクロスオーバーフィルタ、ロールオフフィルタ及び遅延パラメータをビットストリームに送り、原信号のステレオ特性に似せることをもっと可能にすることができる。ステレオ信号を生成するために残響装置を使用する場合、音の最後以後の残響減衰は時には望ましくないかもしれない。しかし、これらの望ましくない残響尾部は、残響信号の利得を変更するだけで、容易に、減衰させるか又は完全に除去することができる。その目的のために、音の終わりを見出す検出器を使用することができる。残響装置が、ある特定の信号、例えば過渡信号でアーチファクトを発生させる場合、それを減衰するためにこれらの信号の検出器を使用することもできる。

【0015】

本発明にかかるステレオ特性を検出する他の方法を以下に述べる。再びL及びRが左側入力信号及び右側入力信号を表わすことにする。次いで対応する信号パワーを $P_L \sim L^2$ 及び $P_R \sim R^2$ によって示す。今、ステレオバランスの大きさは2つの信号パワーの商として、より詳しくは $B = (P_L + e) / (P_R + e)$ として計算することができる。ここでeは、零による除算を排除する、任意の非常に小さい数である。バランスパラメータBはdB単位で表わすことができ、関係 $B_{dB} = 10 \log_{10}(B)$ によって与えられる。一例として、 $P_L = 10 P_R$ 、 $P_L = P_R$ 、及び $P_L = 0.1 P_R$ の3つの場合は、+10dB、0dB、及び-10dBにそれぞれ対応する。明らかに、これらの値は位置「左」、「中央」、及び「右」を意味する。実験から、バランスパラメータのスパンを例えば+/-40dBに制限できることが分かっている。これらの極値は、すでに音が完全に2つのスピーカ又はヘッドホンドライバのうちの1つから発しているように知覚されるためである。この制限は、伝送中に対象とする信号量を低減し、したがってビットレートの低減をもたらす。さらに、順次量子化方式を使用することができ、それによって零付近では小さい量子化ステップを使用し、外側の極値に向かってより大きいステップを使用し、ビットレートをさらに低減する。バランスは長期の伝送に対して長時間一定であることが多い。したがって、必要な平均ビット数を著しく低減する最後の手段を講じることができる。初期バランス値の伝送後、連続バランス値間の差だけを伝送して、エントロピ符号化を使用する。たいていの場合、この差は零であって、したがって可能な限り最も短かいコードワードによって送信される。ビット誤りが起こり得る用途では、制御できない誤り伝播を排除するために、このデルタ符号化は適切な時間間隔でリセットしなければならないことは明らかである。

【0016】

バランスパラメータの最も基本的な復号器の利用法は、両方の出力にモノラル信号を供給し、図2cのブロック227及び229に示すように制御信号Bにより、それに相応し

10

20

30

40

50

て利得を調整することによって、2つの再生チャンネルのいずれかの方向にモノラル信号を単純に偏位させることである。これは、ミキシングデスク上で「パノラマ」ノブを回し、2つのステレオスピーカの間でモノラル信号を合成的に「動かす」ことに似ている。

【0017】

上述した幅パラメータに加えてバランスパラメータを送信して、制御された方法によるサウンドステージにおける音の像の配置と拡散の両方を可能にし、原ステレオ感覚に似せるときに柔軟性を与えることができる。前節で述べた疑似ステレオ生成とパラメータ制御バランスを組み合わせることににおける1つの問題は、中心位置から遠いバランス位置での疑似ステレオ生成器からの望ましくない信号寄与である。これは、モノラル向きの関数をステレオ幅値に適用することによって解決され、結果的に、端側位置のバランス位置ではステレオ幅値の減衰が大きくなり、中心位置に近いバランス位置では減衰が少ないか全くないようになる。

【0018】

これまで説明した方法は、非常に低いビットレートの用途向けに意図されている。より高いビットレートが利用可能な用途では、上記の幅及びバランス法のより精巧な形態を使用することが可能である。幾つかの周波数帯域でステレオ幅検出を行なうことができ、結果的に各周波数帯域について個別のステレオ幅値が得られる。同様にバランス計算をマルチバンド方式で操作することができ、それはモノラル信号が供給される2つのチャンネルに異なるフィルタ曲線を適用することと同等である。図3は、ブロック307、317、及び327によって表わされる図2bに係る1組のN個の疑似ステレオ生成器を、図2cで説明したブロック309、319、及び329で表わされるマルチバンドバランス調整と組み合わせて使用する、パラメトリックステレオ復号器の一例を示す。個々の通過域は、モノラル入力信号Mを1組の帯域フィルタ305、315及び325に供給することによって得られる。バランス調整器からの通過域ステレオ出力は加算器311、321、313、323で加算され、ステレオ出力信号L及びRが形成される。以前のスカラの幅及びバランスパラメータは、ここでは、配列 $W(k)$ 及び $B(k)$ に置き換えられる。図3で、全ての疑似ステレオ生成器及びバランス調整器は独自のステレオパラメータを持つ。しかし、伝送または格納されるデータの総量を低減するために、幾つかの周波数帯域からのパラメータを符号器でグループ毎に平均し、このより少数のパラメータを復号器で対応するグループの幅及びバランスブロックに配置することができる。配列 $W(k)$ 及び $B(k)$ に様々なグループ分け方式及び長さを使用できることは明らかである。 $S(k)$ は幅ブロックの遅延信号経路の利得を表わし、 $D(k)$ は遅延パラメータを表わす。ここでも、 $S(k)$ 及び $D(k)$ はビットストリームでは使用は任意である。

【0019】

パラメトリックバランス符号化法は、特に低周波数帯域の場合、周波数分解能の欠如のため、あるいは異なるバランス位置ではあるが1つの周波数帯域で同時に発生する音声が多すぎるために、多少不安定な挙動をもたらすことがあり得る。これらのバランスの不調は通常ごく短期間の逸脱したバランス値によって特徴付けられ、典型的なものは更新レートに応じた1個又は数個の連続計算値である。邪魔なバランス不調を防止するために、バランスデータに安定化プロセスを適用することができる。このプロセスは、現行時間位置の前後の多数のバランス値を使用して、それらの中央値を算出することができる。その中央値は、現行バランス値のリミッタ値として使用することができ、すなわち、現行バランス値はその中央値を超えることができないものとするすることができる。その結果、現行値は最後の値と中央値との間の範囲によって制限される。必要に応じて、現行バランス値は特定のオーバーシュート係数だけ制限値を超えられるようにすることができる。さらに、中央値の算出に使用されるバランス値の個数だけでなくオーバーシュート係数も周波数従属特性であり、したがって各周波数帯域に対して固有であるとみるべきである。

【0020】

バランス情報の更新率が低い場合、時間分解能が欠如して、ステレオ像の動きと実際の音声との間の同期化に不良を発生させ得る。同期化に関するこの挙動を改善するために、

10

20

30

40

50

音声の識別に基づく補間方式を使用することができる。補間とはここでは、2つの時間的に連続したバランス値間の補間を指す。受信機側でモノラル信号を調べることによって、異なる音声の始まりと終わりに関する情報を得ることができる。1つの方法は、特定の周波数帯域における信号エネルギーの突然の増加又は減少を検出することである。補間は、時間的には、そのエネルギー包絡線からの誘導の後、バランス位置の変更を、好ましくは信号エネルギーをほとんど含まない時間セグメント中に実行するようにすべきである。人間の耳は音の後縁部より入口部に敏感なので、補間方式は、例えばピークホールドをエネルギーに適用することによって、音の開始を見出すことが有利であり、次いでバランス値の増分をピークホールドエネルギーの関数とする。エネルギー値が小さければ増分は大きく、その逆もしかりである。時間的に均等に分散されたエネルギーを含む時間セグメントの場合、すなわち10
幾つかの静止信号の場合など、この補間法は2つのバランス値間の線形補間に等しい。バランス値が左側及び右側のエネルギーの商である場合、左右の対称性のため、対数バランス値が好ましい。対数領域に全補間アルゴリズムを適用する別の利点は、レベルを対数尺に関連付ける人間の耳の傾向である。

【0021】

また、ステレオ幅の利得値の更新率が低い場合、補間をそれに適用することができる。簡単な方法は、時間的に連続する2つのステレオ幅値の間で線形的に補間するものである。幾つかのステレオ幅パラメータを含むより長い時間セグメントに亘ってステレオ幅の利得値を平滑化することによって、ステレオ幅のより安定な挙動を達成することができる。20
様々なアタック時定数及びリリース時定数による平滑化を利用することによって、混合又は差し挟まれた会話及び音楽を含むプログラム素材によく適したシステムが達成される。そのような平滑化フィルタの適切な設計は、短いアタック時定数を使用して短い立上り時間、及びしたがってステレオでの音楽開始への即時対応が得られ、かつ長いリリース時間を使用して長い立下り時間が得られるように行なわれる。突然に会話に入るのに望ましい場合があるワイドステレオモードからモノラルへの高速切換えができるようにするために、この事象を送信することによって、平滑化フィルタを迂回又はリセットすることができる。さらに、アタック時定数、リリース時定数、及び他の平滑化フィルタ特性をも符号器によって信号として送ることができる。

【0022】

疑似音響コーデックからのマスクされたひずみを含む信号の場合、符号化モノラル信号に基づくステレオ情報の導入における1つの共通の問題は、ひずみの脱マスクング効果である。通常「ステレオ脱マスクング」と呼ばれるこの現象は、マスクング基準を満たさない非中心合わせ音声の結果である。ステレオ脱マスクングの問題は、そのような状況を目30
的とした検出器を復号器側に導入することによって、解決又は部分的に解決することができる。信号対マスク比を測定するための公知の技術を使用して、潜在的なステレオ脱マスクングを検出することができる。ひとたび検出されると、それを明示的に送信することができ、あるいはステレオパラメータを簡単に低減させることができる。

【0023】

符号器側において、本発明で教示する1つの選択肢は、入力信号に対しヒルベルト変換器を使用することである。すなわち、2つのチャンネル間に90度移相を導入する。その後40
に2つの信号の加算によってモノラル信号を形成すると、ヒルベルト変換は中心情報に3dBの減衰を導入するので、中心にパンされたモノラル信号と「真」のステレオ信号との間のよりよいバランスが達成される。実際、これは、リードボーカル及びベースギターが一般に単一モノラルソースを使用して録音される、例えば現代ポップ音楽のモノラル符号化を改善する。

【0024】

マルチバンドバランスパラメータ法は、図1に記載した適用の型に限定されず、ステレオ信号のパワースペクトル包絡線を効率的に符号化することが目的である場合はいつでも、有利に使用することができる。したがって、ステレオスペクトル包絡線に加えて対応するステレオ残留が符号化されるステレオコーデックにおけるツールとして使用することが50

できる。全パワー P が $P = P_L + P_R$ によって定義されるとする。ここで P_L 及び P_R は上述した信号パワーである。この定義は、左側と右側の位相関係を考慮していないことに注意されたい。(例えば符号が逆である同一の左側信号と右側信号であっても全パワーは零にならない)。 B と同様に、 P は dB 単位で $P_{dB} = 10 \log_{10} (P / P_{ref})$ と表わすことができる。ここで P_{ref} は任意の基準パワーであり、デルタ値は符号化されたエントロピである。バランスの場合とは対照的に、 P に対しては順次量子化は使用しない。ステレオ信号のスペクトル包絡線を表わすために、 P と B は、必ずしもそうとは限らないが、典型的には人間の聴覚の臨界帯域に関連する帯域幅をもつ1組の周波数帯域に対して計算される。例えば、これらの帯域は、一定帯域幅フィルタバンクにチャンネルをグループ分けすることによって形成することができ、それにより P_L 及び P_R はそれぞれの帯域及び周期時間に対応するサブバンドサンプルの二乗の時間及び周波数平均として計算される。集合 $P_0, P_1, P_2, \dots, P_{N-1}$ 及び $B_0, B_1, B_2, \dots, B_{N-1}$ (ここで下付数字は N 個の帯域表現における周波数帯域を表わす) はデルタ符号化及びハフマン符号化され、伝送又は格納され、最終的に、符号器で計算された量子化値に復号化される。最後のステップは、 P 及び B を P_L 及び P_R に戻すことである。 P 及び B の定義から容易に分かるように、逆の関係は (B の定義における e を無視したとき) $P_L = B P / (B + 1)$ 及び $P_R = P / (B + 1)$ である。

【0025】

上記の包絡線符号化法の1つの特に興味深い用途は、HFRに基づくコーデックのための高帯域のスペクトル包絡線の符号化である。この場合、高帯域残留信号は伝送されない。代わりにこの残留は低帯域から導出される。したがって、残留及び包絡線表現間に厳密な関係はなく、包絡線量子化はより重要である。量子化の効果を調べるために、 P_q 及び B_q が P 及び B の量子化値をそれぞれ表わすものとする。次いで P_q 及び B_q を上記の関係に挿入し、和を形成する。すなわち、

$$P_L q + P_R q = B_q P_q / (B_q + 1) + P_q / (B_q + 1) = P_q (B_q + 1) / (B_q + 1) = P_q$$

となる。ここで興味深い特徴は、 B_q が除去され、全パワーの誤差が P の量子化誤差だけによって決定されることである。これは、たとえ B が密に量子化されても、 P の量子化で十分な精度が使用されるならば、知覚されるレベルが正確であることを暗示している。言い換えると、 B のひずみは、レベルではなく、むしろ空間のひずみに置き換えられる。音源が長時間空間的に静止している限り、ステレオ配合のこのひずみも静止しており、気付きにくい。人間の聴覚の特性のため、中心線に対する角度が大きい場合、 dB の任意の誤差は知覚される角度の小さい誤差に対応するので、すでに述べた通り、ステレオバランスの量子化は外側端に向かって粗くすることができる。

【0026】

周波数従属データ、例えばマルチバンドステレオ幅利得値又はマルチバンドバランス値を量子化する場合、量子化方法の分解能及び範囲は、知覚スケールの特性に一致するように有利に選択することができる。そのようなスケールを周波数に依存させる場合、異なる周波数帯域に対して異なる量子化方法又はいわゆる量子化クラスを選択することができる。異なる周波数帯域を表わす符号化されたパラメータ値は、次いで、場合によっては、たとえ同一値であっても異なるように解釈し、すなわち異なる値に復号化しなければならない。

【0027】

極端信号によりよく対処するために、切換え $L/R - S/D$ 符号化方式と同様に、 P 信号と B 信号は P_L 信号と P_R 信号に適応的に切り換えることができる。 [P C T / S E 0 0 / 0 0 1 5 8] によって教示される通り、包絡線サンプルのデルタ符号化は、特定の瞬間にビット数の点からどちらが最も効率的であるかによって、時間デルタから周波数デルタに切り換えることができる。バランスパラメータもまた、この方式を利用することができる。すなわち、例えばステレオフィールド内を時間の経過と共に移動するソースを考える。明らかに、これは時間の経過に対するバランス値の連続変化に対応し、それはソースの速度対パラメータの更新率によって、大きい時間デルタ値に対応することがあり、エント

ロピ符号化を使用する場合の大きいコードワードに対応する。しかし、ソースが周波数に対して均等な音の放射をもつと仮定すると、バランスパラメータの周波数デルタ値は全ての時間点で零であり、再び小さいコードワードに対応する。したがって、この場合、周波数デルタ符号化を使用すると低いビットレートが達成される。別の例は、ソースが室内で静止しているが、均一でない放射をもつ場合である。この場合は、周波数デルタ値は大きく、時間デルタが好適な選択肢となる。

【 0 0 2 8 】

P / B 符号化方式は、スケーラブルな H F R コーデックを作成する可能性を与える。図 4 を参照されたい。スケーラブルなコーデックは、ビットストリームが 2 つ又はそれ以上の部分に分割されることによって特徴付けられ、高次の部分の受信及び復号化の使用は任意である。例として、以下で一次ビットストリーム 4 1 9 及び二次ビットストリーム 4 1 7 と呼ぶ 2 つのビットストリーム部分を取り上げるが、より多数の部分に拡張が可能であることは明らかである。符号化側 (図 4 a) は、ステレオ入力信号 I N に対して動作する任意のステレオ低帯域符号器 4 0 3 (A D または D A 変換の自明のステップは図示されていない) と、高帯域スペクトル包絡線及び必要に応じて追加のステレオパラメータを推定し、またステレオ入力信号に対して動作するパラメトリックステレオ符号器 4 0 1 と、それぞれ一次及び二次ビットストリーム用の 2 つのマルチプレクサ 4 1 5 及び 4 1 3 とを備えている。この適用では、高帯域包絡線符号化は P / B 動作に固定され、P 信号 4 0 7 はマルチプレクサ 4 1 5 によって一次ビットストリームに送られる一方、B 信号 4 0 5 はマルチプレクサ 4 1 3 によって二次ビットストリームに送られる。

【 0 0 2 9 】

低帯域コーデックの場合、異なる可能性が存在する。すなわち、それは常に S / D モードで動作することができ、S 信号及び D 信号がそれぞれ一次ビットストリーム及び二次ビットストリームに送られる。この場合、一次ビットストリームを復号化すると全帯域のモノラル信号が得られる。言うまでもなく、このモノラル信号は、本発明に係るパラメトリックステレオ方法によって強化することができ、その場合ステレオパラメータも一次ビットストリーム内に配置しなければならない。別の可能性は、ステレオ符号化低帯域信号を、必要に応じて高帯域の幅パラメータ及びバランスパラメータと共に、一次ビットストリームに供給することである。低帯域のステレオ特性は高周波再構成に反映されるので、この場合は、一次ビットストリームを復号化すると、低帯域の真のステレオ及び高帯域の非常に現実的な疑似ステレオが得られる。別の言い方をすると、利用可能な高帯域包絡線表現又はスペクトル粗構造がモノラルであったとしても、合成される高帯域残留又はスペクトル微細構造はそうではない。この型の実施では、二次ビットストリームはより多くの低帯域情報を含むことができ、それは一次ビットストリームのそれと組み合わせられたときに、より高品質の低帯域再生を生み出す。マルチプレクサ 4 1 5 及び 4 1 7 にそれぞれ接続される一次及び二次の低帯域符号器出力信号 4 1 1 及び 4 0 9 は、上述した信号型のどちらでも含むことができるので、図 4 のトポロジーは両方の場合を表わす。

【 0 0 3 0 】

ビットストリームは伝送又は格納され、一次ビットストリーム 4 1 9 だけ、または一次ビットストリーム 4 1 9 及び二次ビットストリーム 4 1 7 の両方が復号器 (図 4 b) に供給される。一次ビットストリームはデマルチプレクサ 4 2 3 によって逆多重化されて、低帯域コア復号器一次信号 4 2 9 及び P 信号 4 3 1 になる。同様に、二次ビットストリームはデマルチプレクサ 4 2 1 によって逆多重化され、低帯域コア復号器二次信号 4 2 7 及び B 信号 4 2 5 になる。低帯域信号は出力 4 3 5 を生成する低帯域復号器 4 3 3 に送られる。出力 4 3 5 は、一次ビットストリームの復号化の場合にのみ、再び上述したどちらの型 (モノラルまたはステレオ) であってもよい。信号 4 3 5 は H F R 装置 4 3 7 に送られ、そこで合成高帯域が生成され、同じく H F R 装置に接続される P 信号に応じて調整される。復号化された低帯域は H F R 装置で高帯域と結合され、低帯域及び / 又は高帯域は、最終的にシステム出力に送られて出力信号 O U T を形成する前に、必要に応じて疑似ステレオ生成器 (これも H F R 装置内に位置する) によって強化される。二次ビットストリーム

4 1 7 が存在する場合、H F R 装置は入力信号 4 2 5 として B 信号をも得、信号 4 3 5 はステレオ信号であり、それによってシステムはフルステレオ出力信号を生成し、疑似ステレオ生成器が存在する場合、それは迂回される。

【 0 0 3 1 】

言い換えると、入力信号のステレオ特性を符号化するための方法は、符号器で前記入力信号のステレオ幅を示す幅パラメータを計算するステップと、復号器で出力信号のステレオ幅を制御するために前記幅パラメータを使用してステレオ出力信号を生成するステップとを含んでいる。その方法は前記符号器で前記入力信号からモノラル信号を形成することをさらに含み、前記復号器での前記ステレオ出力信号の生成は前記モノラル信号に対して作用する疑似ステレオ方法を含む。その方法は前記モノラル信号を 2 つの信号に分割すると共に、前記モノラル信号の遅延バージョンを、前記幅パラメータによって制御されたレベルで、前記 2 つの信号に加えることをさらに含む。その方法は、前記遅延バージョンが前記 2 つの信号に加えられる前に高域フィルタにかけられ、より高い周波数で漸次減衰させられることをさらに含む。その方法は、前記幅パラメータがベクトルであり、前記ベクトルの要素が別個の周波数帯域に対応することをさらに含む。その方法は、前記入力信号がデュアルモノラル型である場合、前記出力信号もまたデュアルモノラル型であることをさらに含む。

10

【 0 0 3 2 】

入力信号のステレオ特性を符号化するための方法は、符号器で前記入力信号のステレオバランスを示すバランスパラメータを計算することを含み、復号器で出力信号のステレオ

20

【 0 0 3 3 】

この方法において、前記符号器で前記入力信号からモノラル信号が形成され、前記復号器でのステレオ出力信号の生成は前記モノラル信号を 2 つの信号に分割することを含み、前記ステレオバランスの制御は前記 2 つの信号のレベルの調整を含む。その方法は、前記入力信号の各チャンネルのパワーが計算され、前記パワー間の商から前記バランスパラメータが計算されることをさらに含む。その方法は、前記パワー及び前記バランスパラメータがベクトルであり、ベクトルの各要素が特定の周波数帯域に対応することをさらに含む。その方法は、前記復号器で、前記モノラル信号の対応するパワーの瞬時値が瞬時補間の勾配をどの程度にすべきかを制御するように、前記バランスパラメータの 2 つの時間的連続値の間で補間されることをさらに含む。その方法は、前記補間法が対数値で表わされたバランス値に対して実行されることをさらに含む。その方法は、前記バランスパラメータの値が前のバランス値とメジアンフィルタ又は他のフィルタプロセスによって他のバランス値から抽出されたバランス値との間の範囲に限定されることをさらに含む。ただし、前記範囲はその境界を特定の範囲だけ移動させることによってさらに拡張することができる。その方法は、バランス値に対する前記限定の境界を抽出する前記方法はマルチバンドシステムの場合には周波数依存性をもつものであることをさらに含む。その方法は、前記パワーのベクトル和として追加レベルパラメータが計算されて前記復号器に送られ、それによって前記復号器に前記入力信号のスペクトル包絡線の表現を与えることをさらに含む。その方法は、前記レベルパラメータ及び前記バランスパラメータが前記パワーによって適応して切り換えられることをさらに含む。その方法は、前記スペクトル包絡線が復号器での H F R プロセスを制御するために使用されることをさらに含む。その方法は、前記レベルパラメータがスケラブルな H F R に基づくステレオコーデックの一次ビットストリームに供給され、前記バランスパラメータが前記コーデックの二次ビットストリームに供給されることをさらに含む。前記モノラル信号及び前記幅パラメータは前記一次ビットストリームに供給される。さらに、前記幅パラメータは、中心位置からより遠いバランス位置に対応するバランス値にはより小さい値を付与する関数によって処理される。その方法は、前記バランスパラメータの量子化が、中心位置付近ではより小さい量子化ステップを使用し、外側の位置に向かってより大きいステップを使用することをさらに含む。その方法は

30

40

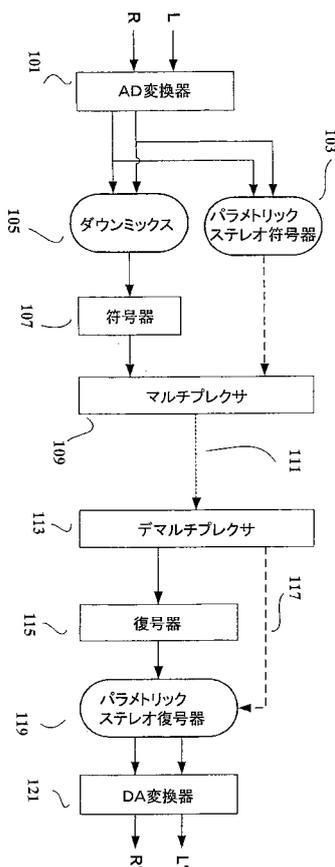
50

、マルチバンドシステムでは周波数依存性をもつ量子化方法を用いて、前記幅パラメータ及び前記バランスパラメータが分解能と範囲に関して量子化されることをさらに含む。その方法は、前記バランスパラメータが時間又は周波数のいずれかで適応してデルタ符号化されることをさらに含む。その方法は、前記モノラル信号を形成する前に前記入力信号がヒルベルト変換器に通されることをさらに含む。

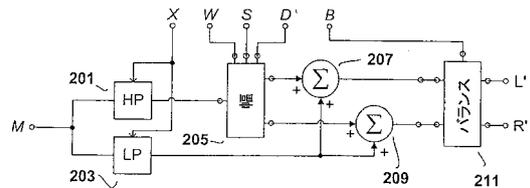
【 0 0 3 4 】

パラメトリックステレオ符号化のための装置は、符号器における、入力信号のステレオ幅を示す幅パラメータを計算するための手段及び前記入力信号からモノラル信号を形成するための手段と、復号器における、出力信号のステレオ幅を制御するために前記幅パラメータを用いて前記モノラル信号からステレオ出力信号を生成するための手段と、を含む。

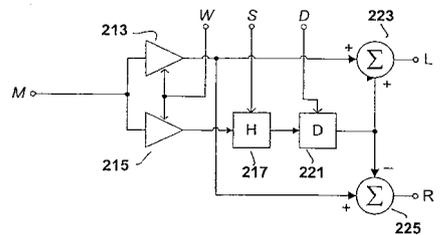
【 図 1 】



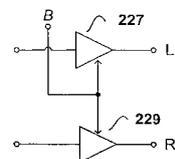
【 図 2 a 】



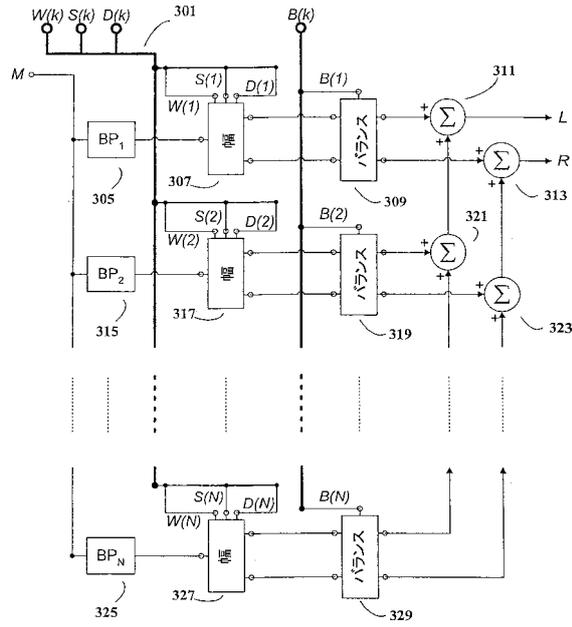
【 図 2 b 】



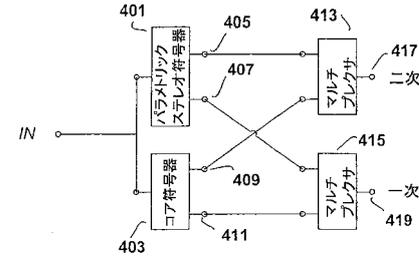
【 図 2 c 】



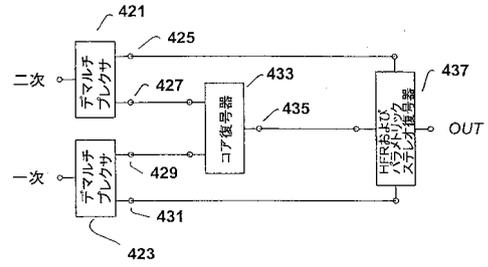
【図3】



【図4a】



【図4b】



フロントページの続き

(31)優先権主張番号 0202159-0

(32)優先日 平成14年7月9日(2002.7.9)

(33)優先権主張国 スウェーデン(SE)

(72)発明者 クリストファー・クジェルリング

スウェーデン国 S - 1 7 0 7 5 ゾルナ ロスティゲン 1 0

(72)発明者 ラルス グスタフ・リルジェリド

スウェーデン国 S - 1 7 1 3 4 ゾルナ ヴィンテルヴェーゲン 1 9

(72)発明者 ヨナス・レーデン

スウェーデン国 S - 1 6 9 3 4 ゾルナ エリック ザンドベルグス ガタ 6

(72)発明者 ヨナス・エンゲデガルド

スウェーデン国 S - 1 1 5 4 3 ストックホルム ヴェンストレムスヴェーゲン 6

審査官 井上 健一

(56)参考文献 特表2001-521648(JP, A)

特表2000-505266(JP, A)

特表平10-504170(JP, A)

特許第4786987(JP, B1)

特許第4878384(JP, B1)

特開2009-217290(JP, A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00

H04S 1/00 - 7/00