

(19) 日本国特許庁(JP)

(12) 公 開 特 許 公 報(A)

(11) 特許出願公開番号
特開2010-72602
(P2010-72602A)

(43) 公開日 平成22年4月2日(2010. 4. 2)

(51) Int.Cl.
G 1 O L 19/00 (2006.01)

F I
G 1 O L 19/00 2 3 0

テーマコード (参考)

審査請求 未請求 請求項の数 9 O L (全 39 頁)

(21) 出願番号	特願2008-243419 (P2008-243419)	(71) 出願人	000002897
(22) 出願日	平成20年9月22日 (2008. 9. 22)		大日本印刷株式会社
			東京都新宿区市谷加賀町一丁目1番1号
		(74) 代理人	100111659
			弁理士 金山 聡
		(74) 代理人	100135954
			弁理士 深町 圭子
		(74) 代理人	100119057
			弁理士 伊藤 英生
		(74) 代理人	100122529
			弁理士 藤枿 裕実
		(74) 代理人	100131369
			弁理士 後藤 直樹
		(72) 発明者	茂出木 敏雄
			東京都新宿区市谷加賀町一丁目1番1号
			大日本印刷株式会社内

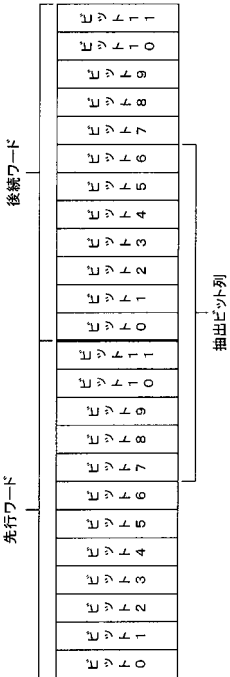
(54) 【発明の名称】 音響信号に対する情報の埋め込み装置および音響信号からの情報の抽出装置

(57) 【要約】

【課題】 前記区切りデータが挿入されていなくても、抽出した所定数のビット値の配列を用いて、本来埋め込まれていた単位ワードのビット配列を誤りなく認識することが可能な音響信号に対する情報の埋め込み装置および音響信号からの情報の抽出装置を提供する。

【解決手段】 埋め込み時に同一ワードを所定回数繰り返し返して埋め込んでおく。抽出時には、先行ワード、後続ワードとして埋め込まれた同一ワード中から、ずれた位置を先頭として抽出ビット列を読み取った場合であっても、あらかじめ埋め込みワードを1ビットずつずらした巡回ビットパターンを記録した巡回符号表を参照し、抽出ビット列と対応する巡回ビットパターンが存在した場合に、対応するビット列を付加情報を構成するビット列として取得する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

時系列のサンプル列で構成される音響信号に対して、 N_w ビットを 1 ワードとして 1 以上のワードから構成される付加情報を聴取不能な状態で繰り返し埋め込む装置であって、

N_w ビットが取り得る 2 の N_w 乗個の全ての基準符号に対して、互いにハミング距離が少なくとも 2 以上となる $N_h (> N_w)$ ビットのハミング符号を割り当てた符号変換テーブルを作成する変換テーブル作成手段と、

前記付加情報中の各ワードについて、前記符号変換テーブルを参照し、対応するハミング符号で構成される N_h ビットのビット配列を作成するビット配列作成手段と、

前記音響信号より、所定数のサンプルを読み込む音響フレーム読込手段と、

前記読み込まれた音響フレームに対して窓関数を用いて周波数変換を行い、スペクトルを得る周波数変換手段と、

前記生成されたスペクトルから所定の周波数範囲の複数のスペクトル集合を抽出し、前記作成されたビット配列に基づいて、当該スペクトル集合の強度を変更する周波数成分変更手段と、

前記変更されたスペクトル集合を含む各スペクトルに対して周波数逆変換を行って、改変音響フレームを生成する周波数逆変換手段と、

前記周波数逆変換された改変音響フレームを順次出力する改変音響フレーム出力手段と、を有し、

前記周波数変換手段、周波数成分変更手段、周波数逆変換手段、改変音響フレーム出力手段は、前記ビット配列作成手段により作成されたビット配列を各ワードについて、所定回数繰り返し実行するものであることを特徴とする音響信号に対する情報の埋め込み装置。

【請求項 2】

N_w ビットを 1 ワードとして 1 以上のワードから構成される付加情報を変換し、 N_h ビットを単位としてあらかじめ聴取不能な状態で埋め込まれた情報を、音響信号から抽出する装置であって、

埋め込み前の元の 1 ワードの付加情報に対応する N_w ビットの基準符号と、当該基準符号に対応する $N_h (> N_w)$ ビットのハミング符号と、当該ハミング符号を構成するビット列をそれぞれ 1 ~ $N_h - 1$ ビットづつ巡回させた $N_h - 1$ 個の巡回ビット列とを対応付けて記録した巡回符号表と、

前記音響信号の所定区間をデジタル化して、所定数のサンプルで構成される音響フレームを獲得する音響フレーム獲得手段と、

前記音響フレームに対して窓関数を用いて周波数変換を行い、スペクトルを得る周波数変換手段と、

前記生成されたスペクトルから所定の周波数範囲の複数のスペクトル集合を抽出し、抽出した各スペクトル集合の強度値に基づいて、埋め込まれていた 1 以上のビット列を抽出する符号化手段と、

前記抽出されたビット列の集合が付加情報 1 ワード分に相当する所定のビット数 N_h に達した場合に、当該 1 ワード分のビット列で前記巡回符号表を参照し、対応する N_w ビットの基準符号を元の 1 ワード分の付加情報として抽出する付加情報抽出手段と、

を有することを特徴とする音響信号からの情報の抽出装置。

【請求項 3】

請求項 2 において、

前記巡回符号表における全てのハミング符号および巡回ビット列は、互いにハミング距離が少なくとも 2 以上であることを特徴とする音響信号からの情報の抽出装置。

【請求項 4】

請求項 2 または請求項 3 において、

N_w ビットが取り得る 2 の N_w 乗個の全ての基準符号に対して、互いにハミング距離が少なくとも 3 以上となる $N_h (> N_w)$ ビットのハミング符号を割り当て、各ハミング符

10

20

30

40

50

号を構成するビット列をそれぞれ $1 \sim N_h - 1$ ビットづつ巡回させた $N_h - 1$ 個の巡回ビット列を作成して、前記基準符号、ハミング符号、 $N_h - 1$ 個の巡回ビット列を対応付けて前記巡回符号表を作成する巡回符号表作成手段をさらに有することを特徴とする音響信号からの情報の抽出装置。

【請求項 5】

請求項 2 から請求項 4 のいずれかにおいて、

前記周波数変換手段は、前記読み込まれた音響フレームのうち、奇数番目、偶数番目の一方を A タイプ、他方を B タイプとし、前記 A タイプの音響フレームに対して第 1 窓関数を用いて周波数変換を行い、前記 B タイプの音響フレームに対して第 2 窓関数を用いて周波数変換を行い、前記第 1 窓関数に対応するスペクトルである第 1 窓スペクトル、前記第 2 窓関数に対応するスペクトルである第 2 窓スペクトルを得るものであり、

10

前記符号化手段は、前記生成された各窓スペクトルから抽出したスペクトル集合の強度値に対して、平均強度レベルを算出し、当該平均強度レベルを規定レベルに一致させるように変換倍率を算出し、前記各スペクトル集合の強度に前記変換倍率を乗じて振幅変換を施すとともに、前記振幅変換を施した各スペクトル集合の強度値に基づいて、埋め込まれていた 1 以上のビット列を抽出するものであることを特徴とする音響信号からの情報の抽出装置。

【請求項 6】

請求項 2 から請求項 5 のいずれかにおいて、

前記符号化手段は、生成された各窓スペクトルから所定の周波数範囲における互いに重複しない 2 つのスペクトル集合を抽出し、前記第 1 窓スペクトル、第 2 窓スペクトルから抽出した高周波側のスペクトル集合の強度値を E_3 、 E_4 、低周波側のスペクトル集合の強度値を E_1 、 E_2 とした場合に、 N_u の整数倍前の対応する音響フレームにおいても同様に、 $E_1 \sim E_4$ に対応するスペクトル集合の強度値を算出し、 N_u の整数倍前の対応する音響フレームについての強度値を、過去のものほど小さくなる 1 未満の重みを乗じた値を累積加算した値を、 $E_1 \sim E_4$ に置換し、置換後の E_1 と E_4 の積と、 E_2 と E_3 の積との大小関係に基づいて、埋め込まれていたビット値を抽出するものであることを特徴とする音響信号からの情報の抽出装置。

20

【請求項 7】

請求項 6 において、

30

前記符号化手段は、1 未満の値を持つ係数 q を用いて、前記スペクトル集合の強度値 E_2 に対しては E_1 を、 E_4 に対しては E_3 を、 E_1 と E_3 に対しては直前の音響フレームにおける各々 E_2 と E_4 に q を乗じた値を減じることにより補正して、強度値 E_1' 、 E_2' 、 E_3' 、 E_4' を得て、

N_u の整数倍前の対応する音響フレームにおいても同様に、 $E_1' \sim E_4'$ に対応するスペクトル集合の強度値を算出し、 N_u の整数倍前の対応する音響フレームについての強度値を、過去のものほど小さくなる 1 未満の重みを乗じた値を累積加算した値を、 $E_1' \sim E_4'$ に置換し、置換後の E_1' と E_4' の積と、 E_2' と E_3' の積との大小関係に基づいて補正ビット値を抽出し、前記 E_1 と E_4 の積と、 E_2 と E_3 の積との大小関係に基づいて抽出されたビット値と、補正ビット値のいずれか一方のビット値を選択するものであることを特徴とする音響信号からの情報の抽出装置。

40

【請求項 8】

請求項 2 から請求項 7 のいずれかにおいて、

前記音響フレーム獲得手段は、前記音響信号から、所定数のサンプルで構成される音響フレームを基準フレームとして獲得する基準フレーム獲得手段と、前記基準フレームと所定サンプルずつ移動させることにより位相を変更して設定される複数の音響フレームを位相変更フレームとして設定する位相変更フレーム設定手段により構成され、

前記符号化手段は、前記抽出したスペクトル集合に基づいて符号判定パラメータを算出するとともに、前記スペクトル集合を、当該スペクトル集合に対応する直前の同タイプの音響フレームにおけるスペクトル集合を利用して補正した補正スペクトル集合に基づいて

50

補正符号判定パラメータを算出する符号判定パラメータ算出手段と、基準フレームが異なる過去の同位相の音響フレームにおいて算出された符号判定パラメータおよび補正符号判定パラメータのいずれかを選択し、当該選択した方の符号判定パラメータに基づいて、前記基準フレームおよび複数の位相変更フレームのうち1つの音響フレームを位相が最適なものであると判断し、当該最適な位相の音響フレームについて判断された前記符号判定パラメータに基づいて、所定の符号を出力する符号出力手段を有するものであることを特徴とする音響信号からの情報の抽出装置。

【請求項9】

コンピュータを、請求項2から請求項8のいずれかに記載の音響信号からの情報の抽出装置として、機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、CD・DVD等を用いた民生・業務用途における鑑賞用のパッケージ音楽分野、放送事業者等が商業目的で配信する放送・ネットワーク音楽配信分野における音楽著作権の保護（不正コピーの監視）および音楽属性情報の提供（楽曲タイトル検索サービス）分野、ミュージアム、イベント会場における展示説明ナレーションに連動した文字情報の提供サービス分野、放送番組やCD/DVDパッケージの音声信号からURLなどの情報を抽出し、携帯電話を用いて所定のコンテンツに関連するwebサイトにアクセスして詳細情報を抽出したり、アンケートに回答したりする非接触なインターネットのゲートウェイサービス分野に関する。

【背景技術】

【0002】

従来、流れている音楽のタイトル等を知ることができる楽曲属性情報の提供サービスとして、放送された音楽に対して日時と地域を放送局に照会したり、携帯電話で流れている音楽断片を録音してデータベースに登録されているメロディーと照合したりするサービスが行われている。本出願人は、これをさらに発展させ、音響信号の周波数成分の比率を属性情報のビット値に応じて変更することにより、属性情報（付加情報）を埋め込む手法を提案している（特許文献1参照）。

【0003】

【特許文献1】特開2006-323246号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

上記特許文献1に記載の発明では、埋め込むビットデータの0または1の2値に対応した符号に加え、1種類の区切りデータを設けた2値構成になっており、この区切りデータは所定数のビットデータで構成されるワードデータ間（特許文献1ではバイトデータ間）、または埋め込み対象の音響信号のレベルが埋め込み可能な状態でない場合に適宜挿入される。このようにして、属性情報が埋め込まれた音響信号を音として再生した状態で、その音を録音して取得し、解析して埋め込まれていた属性情報を抽出する際、再生・録音状態によっては、前記区切りデータを適切に取り込むことができず、隣接するワードデータのビットとまたがり、ビットデータがずれた状態でワードデータが抽出されてしまい易いという問題がある。すなわち、埋め込まれていた本来のビット値を適切に抽出することができても、隣接するワードデータ間との区切りの判定を誤るだけで、大幅な文字化けを発生させてしまうという問題がある。埋め込まれる属性情報は、通常、1ワードの複数ビットで所定の意味を持つように構成されているため、1ビットでもずれると、意味が全く変わってしまうことがある。

【0005】

そこで、本発明は、前記区切りデータが挿入されていなくても、抽出した所定数のビット値の配列を用いて、本来埋め込まれていた単位ワードのビット配列を誤りなく認識する

10

20

30

40

50

ことが可能な音響信号に対する情報の埋め込み装置および音響信号からの情報の抽出装置を提供することを課題とする。

【課題を解決するための手段】

【0006】

上記課題を解決するため、本発明では、時系列のサンプル列で構成される音響信号に対して、 N_w ビットを1ワードとして1以上のワードから構成される付加情報を聴取不能な状態で繰り返し埋め込む装置であって、 N_w ビットが取り得る2の N_w 乗個の全ての基準符号に対して、互いにハミング距離が少なくとも2以上となる N_h ($> N_w$) ビットのハミング符号を割り当てた符号変換テーブルを作成する変換テーブル作成手段と、前記付加情報中の各ワードについて、前記符号変換テーブルを参照し、対応するハミング符号で構成される N_h ビットのビット配列を作成するビット配列作成手段と、前記音響信号より、所定数のサンプルを読み込む音響フレーム読込手段と、前記読み込まれた音響フレームに対して窓関数を用いて周波数変換を行い、スペクトルを得る周波数変換手段と、前記生成されたスペクトルから所定の周波数範囲の複数のスペクトル集合を抽出し、前記作成されたビット配列に基づいて、当該スペクトル集合の強度を変更する周波数成分変更手段と、前記変更されたスペクトル集合を含む各スペクトルに対して周波数逆変換を行って、改変音響フレームを生成する周波数逆変換手段と、前記周波数逆変換された改変音響フレームを順次出力する改変音響フレーム出力手段と、を有し、前記周波数変換手段、周波数成分変更手段、周波数逆変換手段、改変音響フレーム出力手段は、前記ビット配列作成手段により作成されたビット配列を各ワードについて、所定回数繰り返して実行するものである音響信号に対する情報の埋め込み装置を提供する。

10

20

【0007】

また、本発明では、 N_w ビットを1ワードとして1以上のワードから構成される付加情報を変換し、 N_h ビットを単位としてあらかじめ聴取不能な状態で埋め込まれた情報を、音響信号から抽出する装置であって、埋め込み前の元の1ワードの付加情報に対応する N_w ビットの基準符号と、当該基準符号に対応する N_h ($> N_w$) ビットのハミング符号と、当該ハミング符号を構成するビット列をそれぞれ1~ $N_h - 1$ ビットずつ巡回させた $N_h - 1$ 個の巡回ビット列とを対応付けて記録した巡回符号表と、前記音響信号の所定区間をデジタル化して、所定数のサンプルで構成される音響フレームを獲得する音響フレーム獲得手段と、前記音響フレームに対して窓関数を用いて周波数変換を行い、スペクトルを得る周波数変換手段と、前記生成されたスペクトルから所定の周波数範囲の複数のスペクトル集合を抽出し、抽出した各スペクトル集合の強度値に基づいて、埋め込まれていた1以上のビット列を抽出する符号化手段と、前記抽出されたビット列の集合が付加情報1ワード分に相当する所定のビット数 N_h に達した場合に、当該1ワード分のビット列で前記巡回符号表を参照し、対応する N_w ビットの基準符号を元の1ワード分の付加情報として抽出する付加情報抽出手段を有する音響信号からの情報の抽出装置を提供する。

30

【発明の効果】

【0008】

本発明によれば、音響信号に1以上のワードで構成される付加情報を埋め込むにあたり、各ワードを所定回数繰り返して埋め込み、抽出側では、抽出した1ワード分のビット列で、埋め込み時の付加情報に対応するビット列とそのビット列を1ビットずつずらしたビット列とを対応付けた巡回符号表を参照し、埋め込み時に相当するビット列を抽出するようにしたので、埋め込み時のワード間の区切り位置が未知であっても、本来埋め込まれていたビット配列とずれた状態であっても、1ワードを誤りなく認識することが可能となる。

40

【発明を実施するための最良の形態】

【0009】

(1. 本発明の基本概念)

本発明では、付加情報を埋め込む際、ワード単位で所定回数繰り返して埋め込み、抽出時には、抽出したワード単位の情報を、1ビットずつずらした巡回ビットパターンと照合

50

するようにしている。例えば、1ワード12ビットとし、1ワードを2回繰り返して埋め込んだ場合の例を図1に示す。この場合、同じものを2回繰り返して埋め込んでいるため、先行ワードと後続ワードは全く同じビット配列である。一方、抽出時には、読み込みの先頭位置を認識することは困難であるため、任意の位置から読み込まれることになる。

【0010】

例えば、図1において先行ワードのビット7から読み込んだとすると、後続ワードのビット6までで12ビットとなる。ここで、先行ワードと後続ワードは全く同じビット配列であるため、抽出ビット列をどちらかにずらせば必ず、埋め込み時の付加情報1ワードと一致することになる。逆に、あらかじめ、付加情報1ワードを1ビットずつずらしたビットパターンを11個用意しておけば、それらのいずれかと抽出ビット列は一致することになる。本発明は、このような考え方を利用したものである。

10

【0011】

(2.1.埋め込み装置の構成)

まず、本発明に係る音響信号に対する情報の埋め込み装置について説明する。図2は、本発明に係る音響信号に対する情報の埋め込み装置の構成を示す機能ブロック図である。図2において、10は音響フレーム読込手段、20は周波数変換手段、30は周波数成分変更手段、40は周波数逆変換手段、50は改変音響フレーム出力手段、60は記憶手段、61は音響信号記憶部、62は付加情報記憶部、63は改変音響信号記憶部、70はビット配列作成手段、80は変換テーブル作成手段である。なお、図2に示す装置は、ステレオ音響信号、モノラル音響信号の両方に対応可能であるが、ここでは、ステレオ音響信号に対して処理を行う場合について説明していく。

20

【0012】

音響フレーム読込手段10は、付加情報の埋め込み対象とする元のステレオ音響信号の各チャンネルから所定数のサンプルを1フレームとして読み込む機能を有している。周波数変換手段20は、音響フレーム読込手段10が読み込んだ音響信号のフレームをフーリエ変換等により周波数変換してスペクトルを生成する機能を有している。周波数成分変更手段30は、生成されたスペクトルから所定の周波数範囲に相当するスペクトル集合を複数抽出し、付加情報記憶部62から抽出した付加情報よりビット配列作成手段70が作成したビット配列に基づいて、スペクトル集合の状態を変更する機能を有している。周波数逆変換手段40は、変更されたスペクトル集合を含む複数のスペクトルに対して周波数逆変換を行うことにより、改変音響フレームを生成する機能を有している。改変音響フレーム出力手段50は、生成された改変音響フレームを順次出力する機能を有している。

30

【0013】

記憶手段60は、付加情報を埋め込む対象とするステレオ音響信号を記憶した音響信号記憶部61と、ビット列として構成され、ステレオ音響信号に埋め込まれる付加情報を記憶した付加情報記憶部62と、付加情報埋め込み後の改変音響信号を記憶する改変音響信号記憶部63を有しており、その他処理に必要な各種情報を記憶するものである。ビット配列作成手段70は、付加情報記憶部62から付加情報を抽出し、付加情報の各ワードについて、変換テーブル作成手段80を参照し、対応するビット配列を作成する機能を有している。変換テーブル作成手段80は、 N_w ビットが取り得る2の N_w 乗個の全ての基準符号に対して、互いにハミング距離が少なくとも2以上となる $N_h (> N_w)$ ビットのハミング符号を割り当てることにより、 N_w ビットの基準符号と N_h ビットのハミング符号が対応付けられた符号変換テーブルを作成する機能を有している。基準符号は、 $N_w = 7$ の場合、10進で表現すると、0~127の値をとり得る。

40

【0014】

なお、付加情報とは、音響情報に付加して埋め込むべき情報であり、タイトルやアーティスト名等の属性情報、および属性情報以外の他の情報を含むものである。本実施形態では、上記 N_w 、 N_h をそれぞれ $N_w = 7$ 、 $N_h = 12$ としている。本実施形態では、付加情報のコード形式としてASCIIコードを採用するため、付加情報においては $N_w = 7$ となり、7ビットが1ワードとなる。また、ビット配列作成手段70が作成するビット配

50

列は、12ビットとなり、ビット配列作成後は、これを1ワードとして処理している。図2に示した各構成手段は、現実にはコンピュータおよびその周辺機器等のハードウェアに専用のプログラムを搭載することにより実現される。すなわち、コンピュータが、専用のプログラムに従って各手段の内容を実行することになる。

【0015】

(2.2.埋め込み装置の処理動作)

次に、図2に示した音響信号に対する情報の埋め込み装置の処理動作について説明する。音響フレーム読込手段10は、音響信号記憶部61に記憶されたステレオ音響信号の左右の各チャンネルから、それぞれ所定数Nのサンプルを1音響フレームとして読み込む。音響フレーム読込手段10が読み込む1音響フレームのサンプル数Nは、適宜設定することができ、サンプリング周波数が44.1kHzの場合、4096サンプル程度とすることが望ましい。したがって、音響フレーム読込手段10は、左チャンネル、右チャンネルについてそれぞれ4096サンプルずつ、順次音響フレームとして読み込んでいくことになる。

10

【0016】

音響フレームとしては、AタイプとBタイプが存在する。Aタイプの音響フレーム、Bタイプの音響フレームは、それぞれ同タイプの先行する音響フレームの最後のサンプルの次のサンプルを先頭サンプルとして設定される。そして、AタイプとBタイプの音響フレームは互いに所定数(本実施形態では2048)のサンプルを重複して設定される。例えば、Aタイプの音響フレームを先頭からA1、A2、A3...とし、Bタイプの音響フレームを先頭からB1、B2、B3...とすると、A1はサンプル1~4096、A2はサンプル4097~8192、A3はサンプル8193~12288、B1はサンプル2049~6144、B2はサンプル6145~10240、B3はサンプル10241~14336となる。なお、AタイプとBタイプは相対的なものであるため、どちらが先であっても良い。すなわち、上記とは逆にA1がサンプル2049~6144、A2がサンプル6145~10240、A3がサンプル10241~14336、B1がサンプル1~4096、B2がサンプル4097~8192、B3がサンプル8193~12288であっても良い。

20

【0017】

周波数変換手段20は、振幅変換後の音響フレームに対して周波数変換を行って、その音響フレームのスペクトルを得る。具体的には、窓関数を利用して周波数変換を行う。周波数変換としては、フーリエ変換、ウェーブレット変換その他公知の種々の手法を用いることができる。本実施形態では、フーリエ変換を用いた場合を例にとって説明する。

30

【0018】

一般に、所定の信号に対してフーリエ変換を行う場合、信号を所定の長さに区切って行う必要があるが、この場合、所定長さの信号に対してそのままフーリエ変換を行うと、擬似高調波成分が発生する。そこで、一般にフーリエ変換を行う場合には、ハニング窓と呼ばれる窓関数を用いて信号の値を変化させた後、変化後の値に対してフーリエ変換を実行する。

40

【0019】

本実施形態においても、窓関数を利用しているが、Aタイプの音響フレーム、Bタイプの音響フレームで、使用する窓関数を分けている。本実施形態では、図3(a)(b)に示したような第1窓関数 $W(1, i)$ 、第2窓関数 $W(2, i)$ を用意し、抽出側で認識し易いようにした。第1窓関数 $W(1, i)$ は、Aタイプの音響フレームに対して用いるためのものであり、図3(a)に示すように所定のサンプル番号*i*の位置において、最大値1をとり、後部においては、最小値0をとるように設定されている。どのサンプル番号の場合に最大値をとるかについては、窓関数 $W(1, i)$ の設計によって異なってくるが、本実施形態では、後述する〔数式1〕で定義される。Aタイプの音響フレームについてのフーリエ変換は、この窓関数 $W(1, i)$ を乗じたものに対して行われることになる。

50

【0020】

また、第2窓関数 $W(2, i)$ は、Bタイプの音響フレームに対して用いるためのものであり、図3(b)に示すように、所定のサンプル番号 i の位置において、最大値1を取り、前部においては、最小値0をとるように設定されている。どのサンプル番号の場合に最大値をとるかについては、窓関数 $W(2, i)$ の設計によって異なってくるが、本実施形態では、後述する〔数式2〕で定義される。Bタイプの音響フレームについてのフーリエ変換は、この窓関数 $W(2, i)$ を乗じたものに対して行われることになる。

【0021】

なお、上述のように、本実施形態においては、音響フレームは重複して読み込まれる。すなわち、奇数番目の音響フレームと偶数番目の音響フレームは、所定数のサンプルを重複して読み込む。上記のように、奇数フレームと偶数フレームでは、用いられる窓関数が異なるが、奇数フレームと偶数フレームは単に奇数か偶数かの違いだけであるため、どちらに対してどちらの処理を行っても良い。したがって、本明細書では、奇数フレーム、偶数フレームの一方をAタイプフレーム、他方をBタイプフレームと呼ぶことにする。本実施形態では、奇数フレームをAタイプフレーム、偶数フレームをBタイプフレームとして説明するが、逆に偶数フレームをAタイプフレーム、奇数フレームをBタイプフレームとしても良い。

【0022】

本実施形態では、窓関数 $W(1, i)$ 、 $W(2, i)$ は、以下の〔数式1〕〔数式2〕で定義される。なお、図3において、横軸は時間軸(i)である。 i は、後述するように、各音響フレーム内の N 個のサンプルに付した通し番号であるため時刻 t に比例している。また、図3(a)(b)において縦軸は信号の振幅値(レベル)を示す。図3(a)(b)において縦軸は窓関数 $W(1, i)$ 、 $W(2, i)$ の値を示しており、 $W(1, i)$ 、 $W(2, i)$ の最大値はいずれも1である。

【0023】

〔数式1〕

$i = N/8$ のとき、 $W(1, i) = 0.0$
 $N/8 < i \leq 3N/8$ のとき、 $W(1, i) = 0.5 - 0.5 \cos(4\pi(i - N/8)/N)$
 $3N/8 < i \leq 11N/16$ のとき、 $W(1, i) = 1.0$
 $11N/16 < i \leq 13N/16$ のとき、 $W(1, i) = 0.5 + 0.5 \cos(8\pi(i - 11N/16)/N)$
 $i > 13N/16$ のとき、 $W(1, i) = 0.0$

【0024】

〔数式2〕

$i = 3N/16$ のとき、 $W(2, i) = 0.0$
 $3N/16 < i \leq 5N/16$ のとき、 $W(2, i) = 0.5 - 0.5 \cos(8\pi(i - 3N/16)/N)$
 $5N/16 < i \leq 5N/8$ のとき、 $W(2, i) = 1.0$
 $5N/8 < i \leq 7N/8$ のとき、 $W(2, i) = 0.5 + 0.5 \cos(4\pi(i - 5N/8)/N)$
 $i > 7N/8$ のとき、 $W(2, i) = 0.0$

【0025】

なお、図3および上記〔数式1〕〔数式2〕から明らかなように、窓関数 $W(1, i)$ と $W(2, i)$ は、互いに非対称な形状である。これは、後述する抽出側において、両者の識別を容易にするためである。

【0026】

本発明においては、奇数フレームと偶数フレームを、所定サンプルずつ重複して読み込むため、情報の埋め込みを行った後、音響信号に復元する際に、窓関数を乗じた奇数フレームと、窓関数を乗じた偶数フレームの重複サンプルを加算した場合に、ほぼ元の値に戻るようしなければならない。このため、奇数フレームと偶数フレームの重複部分におい

10

20

30

40

50

て、窓関数 $W(1, i)$ 、 $W(2, i)$ を加算すると、全区間固定値 1 になるように定義されている。

【0027】

周波数変換手段 20 が、A タイプの音響フレームに対してフーリエ変換を行う場合は、左チャンネル信号 $X_l(i)$ 、右チャンネル信号 $X_r(i)$ ($i = 0, \dots, N-1$) に対して、窓関数 $W(1, i)$ を用いて、以下の〔数式 3〕に従った処理を行い、左チャンネルに対応する変換データの実部 $A_l(1, j)$ 、虚部 $B_l(1, j)$ 、右チャンネルに対応する変換データの実部 $A_r(1, j)$ 、虚部 $B_r(1, j)$ を得る。

【0028】

〔数式 3〕

$$A_l(1, j) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X_l(i) \cdot \cos(2\pi i j / N)$$

$$B_l(1, j) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X_l(i) \cdot \sin(2\pi i j / N)$$

$$A_r(1, j) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X_r(i) \cdot \cos(2\pi i j / N)$$

$$B_r(1, j) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X_r(i) \cdot \sin(2\pi i j / N)$$

【0029】

周波数変換手段 20 が、B タイプの音響フレームに対してフーリエ変換を行う場合は、左チャンネル信号 $X_l(i + N/2)$ 、右チャンネル信号 $X_r(i + N/2)$ ($i = 0, \dots, N-1$) に対して、窓関数 $W(2, i)$ を用いて、以下の〔数式 4〕に従った処理を行い、左チャンネルに対応する変換データの実部 $A_l(2, j)$ 、虚部 $B_l(2, j)$ 、右チャンネルに対応する変換データの実部 $A_r(2, j)$ 、虚部 $B_r(2, j)$ を得る。

【0030】

〔数式 4〕

$$A_l(2, j) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X_l(i + N/2) \cdot \cos(2\pi i j / N)$$

$$B_l(2, j) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X_l(i + N/2) \cdot \sin(2\pi i j / N)$$

$$A_r(2, j) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X_r(i + N/2) \cdot \cos(2\pi i j / N)$$

$$B_r(2, j) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X_r(i + N/2) \cdot \sin(2\pi i j / N)$$

【0031】

上記〔数式 3〕〔数式 4〕において、 i は、各音響フレーム内の N 個のサンプルに付した通し番号であり、 $i = 0, 1, 2, \dots, N-1$ の整数値をとる。また、 j は周波数の値について、値の小さなものから順に付した通し番号であり、 i と同様に $j = 0, 1, 2, \dots, N/2 - 1$ の整数値をとる。サンプリング周波数が 44.1 kHz 、 $N = 4096$ の場合、 j の値が 1 つ異なると、周波数が 10.8 Hz 異なることになる。

【0032】

上記〔数式 3〕〔数式 4〕に従った処理を実行することにより、各音響フレームの各窓関数に対応するスペクトルが得られる。続いて、周波数成分変更手段 30 が、生成されたスペクトルから所定周波数範囲のスペクトル集合を抽出する。本実施形態では、 F_1 以上 F_2 以下の範囲のものを抽出する。

【0033】

周波数変換手段 20 は、周波数変換を実行した後に、各スペクトル集合の平均値が設定値になるように変換する。いわゆる振幅を変換する処理を行う。この振幅変換は、各フレームごとのレベル差を軽減し、適切な情報の埋め込みが可能になるようにすることを目的としているため、設定値としては適宜設定することができるが、本実施形態では、128 に設定するようにしている。

【0034】

周波数変換手段 20 は、得られたスペクトル集合に対して、振幅変換を行うにあたり、まず変換倍率の算出を行う。変換倍率の算出は、所定周波数範囲のスペクトル集合の実効

10

20

30

40

50

強度値の平均である平均実効値で上記設定値を除算することにより行う。具体的には、LチャンネルのAタイプの音響フレーム、Bタイプの音響フレーム、RチャンネルのAタイプの音響フレーム、Bタイプの音響フレームのスペクトル集合についての変換倍率 $Z_l(1)$ 、 $Z_l(2)$ 、 $Z_r(1)$ 、 $Z_r(2)$ は、以下の〔数式5〕に従った処理により算出される。対象音響フレームが無音に近く、分母の二乗総和値が所定の値に満たない場合は、変換倍率を1.0に設定し、振幅変換は行わないようにする。振幅変換が実行されることにより、全ての音響フレームは、元の状態における信号強度に関わらず、各音響フレームのスペクトル集合の平均実効値が設定値となるように振幅変換された状態で、情報の埋め込みが行われることになる。

【0035】

10

〔数式5〕

$$Z_l(1) = Z_o / [\sum_{j=m, \dots, m+2G-1} \{A_l(1, j)^2 + B_l(1, j)^2\}]^{1/2}$$

$$Z_l(2) = Z_o / [\sum_{j=m, \dots, m+2G-1} \{A_l(2, j)^2 + B_l(2, j)^2\}]^{1/2}$$

$$Z_r(1) = Z_o / [\sum_{j=m, \dots, m+2G-1} \{A_r(1, j)^2 + B_r(1, j)^2\}]^{1/2}$$

$$Z_r(2) = Z_o / [\sum_{j=m, \dots, m+2G-1} \{A_r(2, j)^2 + B_r(2, j)^2\}]^{1/2}$$

【0036】

上記〔数式5〕において、 $Z_o = 2G$ であり、本実施形態では、 $G = 80$ であるため、 $Z_o = 160$ である。

【0037】

さらに、 $j = m, \dots, m + 2G - 1$ （周波数 $F_1, \dots, F_2$ に相当）の範囲で、 $A_l(1, j)$ および $B_l(1, j)$ の各々の要素に対して $Z_l(1)$ を乗じ、 $A_l(2, j)$ および $B_l(2, j)$ の各々の要素に対して $Z_l(2)$ を乗じ、 $A_r(1, j)$ および $B_r(1, j)$ の各々の要素に対して $Z_r(1)$ を乗じ、 $A_r(2, j)$ および $B_r(2, j)$ の各々の要素に対して $Z_r(2)$ を乗じることにより、振幅変換を行う。以下説明において、 $A_l(1, j)$ 、 $B_l(1, j)$ 、 $A_l(2, j)$ 、 $B_l(2, j)$ 、 $A_r(1, j)$ 、 $B_r(1, j)$ 、 $A_r(2, j)$ 、 $B_r(2, j)$ はこれらの振幅変換を行った値とする。

20

【0038】

周波数成分変更手段30は、Aタイプの音響フレームについて、ビット配列作成手段70が作成したビット配列に応じて、所定周波数成分の割合を変更する処理を行う。本発明では、ビット配列を1ビットずつ読み込み、Aタイプ、Bタイプの1対の音響フレームに対して1ビットの情報を埋め込む。埋め込まれる1ビットの値は、“0”“1”の2通りがある。本実施形態では、これらを値1、値2と定義する。2種類の符号を埋め込むことができるという点で、これらを符号1、符号2と表現することも可能である。この際、“0”“1”の2通りのうち、いずれを値1、値2（符号1、符号2）と定義しても良い。抽出側において、埋め込み側で埋め込まれた1ビットが特定できれば良いためである。したがって、この定義は、埋め込み側と抽出側で一致している必要がある。

30

【0039】

周波数成分の変更の手法としては、様々なものが考えられるが、本実施形態では、人間の聴覚心理特性である音脈分凝の原理を利用する。ここで、音脈分凝について説明しておく。音脈分凝とは、時系列に高い音と低い音が交互に進行するパターンに対して、人間が、あたかも高低2つのトラックが連続して流れるように音を補間して聞いてしまう錯覚現象である。

40

【0040】

たとえば、図4(a)に示すように、3つの低い音である低音1、低音3、低音5と3つの高い音である高音2、高音4、高音6が、低音1、高音2、低音3、高音4、低音5、高音6の順に演奏されている場合を考えてみる。低音1、低音3、低音5と高音2、高音4、高音6は1オクターブ程度離れており、低い音と高い音は同時に演奏されることはないが、時間的間隔は、ほぼ連続しているものとする。この場合、人間には、図4(b)に示すように、低い音と高い音が演奏されていない部分も補間されて演奏されているよう

50

に聞こえる。すなわち、実際の演奏は単旋律であるのに、人間には、図 4 (b) に示すように、高音 1 ´、高音 3 ´、高音 5 ´、低音 2 ´、低音 4 ´、低音 6 ´ が補間されて複旋律であるように聞こえる。例えば、高音 3 ´ は高音 2 と高音 4 が連続的につながるように、高音 2 と高音 4 の平均的な成分に聞こえる。また、末端部の高音 1 ´ と低音 6 ´ についても、隣接する高音 2 と低音 5 に近い成分で、各々低音 1 と高音 6 に対応するように聞こえる。しかし、マイクロフォン等の電気音響装置は、図 4 (a) に示したものをそのままの音として取得することになる。本実施形態では、このような性質を利用する。なお、補間された音は、前後に演奏されている音と必ずしも同じレベルに補間されて聴こえるのではなく、大雑把に言って、前後に演奏されている音の 50 % 程度があたかも補間されているように聞こえる。

10

【 0 0 4 1 】

具体的には、埋め込み装置で、音響信号から抽出した所定数のサンプルで構成される音響フレーム内の所定周波数範囲の周波数成分を変更するに際し、その強弱が音脈分凝を発生させるような状態に変更する。これにより、人間には、音が途切れたようには聞こえないが、抽出装置では、その明確な変化を認識することができる。

【 0 0 4 2 】

本実施形態では、音脈分凝の原理を利用して、音響フレームの変更対象周波数帯の成分を 2 つの状態に変更し、1 ビットの情報を埋め込むようにしている。ここで、埋め込み処理の前後における音響フレームの所定周波数成分の変化の状態について説明する。図 5 に、本実施形態による A タイプ、B タイプの各チャンネル 1 音響フレームの所定周波数成分の状態を示す。図 5 に示す各音響フレームにおいて、横軸は時間方向、縦軸は周波数方向を示している。

20

【 0 0 4 3 】

図 5 においては、縦軸の周波数方向において、周波数領域が 4 つに区分されているが、上から 2 番目と 3 番目の領域、すなわち、周波数 F 1 以上 F 2 以下の間が変更対象周波数帯であり、最上部すなわち周波数 F 2 超、最下部すなわち F 1 未満は、変更対象でない周波数帯である。すなわち、本実施形態では、周波数 F 1 以上 F 2 以下を所定周波数範囲として、スペクトル集合の強度を変更することになる。図 5 (a) に示すように、L チャンネルの A タイプ音響フレームの変更対象周波数帯については、高周波側におけるスペクトルを L 1 U、低周波側におけるスペクトルを L 1 D で表現し、R チャンネルの A タイプ音響フレームの変更対象周波数帯については、高周波側におけるスペクトルを R 1 U、低周波側におけるスペクトルを R 1 D で表現することとする。また、図 5 (b) に示すように、L チャンネルの B タイプ音響フレームの変更対象周波数帯については、高周波側におけるスペクトルを L 2 U、低周波側におけるスペクトルを L 2 D で表現し、R チャンネルの B タイプ音響フレームの変更対象周波数帯については、高周波側におけるスペクトルを R 2 U、低周波側におけるスペクトルを R 2 D で表現することとする。

30

【 0 0 4 4 】

本実施形態で、符号 1 を埋め込む場合、図 5 (c) (e) に示すように、L 1 D と L 2 U の強度の積、R 1 D と R 2 U の強度の積を相対的に強い状態に変更し、L 1 U と L 2 D の強度の積、R 1 U と R 2 D の強度の積を相対的に弱い状態に変更する。この状態を“状態 1”と呼ぶことにする。符号 2 を埋め込む場合は、図 5 (d) (f) に示すように、L 1 U と L 2 D の強度の積、R 1 U と R 2 D の強度の積を相対的に強い状態に変更し、L 1 D と L 2 U の強度の積、R 1 D と R 2 U の強度の積を相対的に弱い状態に変更する。この状態を“状態 2”と呼ぶことにする。網掛けされた部分の濃さは、同濃度のものは、積を求めるための組となるスペクトルであることを示している。網掛けの色が濃い方が相対的に強度が強い状態に変更される組を示している。

40

【 0 0 4 5 】

本実施形態では、図 5 (c) (e) または (d) (f) に示すような 2 つの状態に A タイプ、B タイプの音響フレームの周波数成分を変更することにより、情報の埋め込みを行っている。2 つの状態であるので 1 ビット分の情報量に相当する。

50

【 0 0 4 6 】

本実施形態では、上記変更対象周波数帯 $F_1 \sim F_2$ を、“ 1 . 7 k H z ~ 3 . 4 k H z ” に設定する。これは、以下のような理由による。すなわち、音声通信として普及度の高い携帯電話機を受信端末として利用する場合、上限については、電話回線帯域および携帯電話機の上限である 3 . 4 k H z とする必要がある。これは、携帯電話機の内部に搭載されているローパスフィルターが、電話交換機に合わせて、3 . 4 k H z 以下に対応したものである。そこで、下限については、上限の 3 . 4 k H z から 1 オクターブ下がった 1 . 7 k H z とすることとした。なお、“ 1 . 7 k H z ” “ 3 . 4 k H z ” という値は、代表的な値であり、必ずしも正確な値である必要はなく、そこから若干ずれた値であっても良い。

10

【 0 0 4 7 】

図 5 に示した例では、相対的に強い状態と弱い状態に変更することについて説明したが、この強弱の程度については、状況に応じて設定することが可能である。両者の割合は、以下に述べるように、大きいほど抽出時の精度が高くなるが、補間される割合は不完全になり、再生時に不連続成分によるノイズが聞こえる。一方、両者の割合が均等であるほど再生品質は原音に近づくが、埋め込んだピットの抽出ができなくなり、再生品質と抽出精度はトレードオフの関係になる。例えば、強い方を 1 0 0 %、弱い方を 0 % とした場合、音脈分凝により、補間される部分の音は図 5 のような変更を行なう前に原音響信号で鳴っていた音の 5 0 % 程度となることが確認されている。

【 0 0 4 8 】

20

そこで、強い方を 7 0 %、弱い方を 3 0 % とした場合、音脈分凝により、補間される部分の音は図 5 のような変更を行なう前に原音響信号で鳴っていた音とほぼ同程度となり、この割合が抽出精度を維持できる限界であることが確認されている。このため、相対的に強いスペクトル集合と、相対的に弱いスペクトル集合の強度の割合を 7 0 % と 3 0 % とすることを目標に設定することが好ましい。これを実現するため、本実施形態では、後述する具体的な処理において、強い状態を設定するための係数 $\alpha = 0 . 7$ 、弱い状態を設定するための係数 $\beta = 0 . 3$ とする。ただし、強い状態に変更すべきスペクトル集合の強度が元々小さい場合には、係数 α 、 β を補正する必要がある。このため、周波数成分変更手段 3 0 は、まず、以下の〔数式 6〕に従った処理を実行することにより、弱い状態に変更すべきスペクトル集合に対する強い状態に変更すべきスペクトル集合の強度割合 y を算出する。

30

【 0 0 4 9 】

〔数式 6〕

$$E1d = \sum_{j=m, \dots, m+G-1} \{ A1(1, j)^2 + B1(1, j)^2 \}$$

$$E2d = \sum_{j=m, \dots, m+G-1} \{ A1(2, j)^2 + B1(2, j)^2 \}$$

$$E1u = \sum_{j=m+G, \dots, m+2G-1} \{ A1(1, j)^2 + B1(1, j)^2 \}$$

$$E2u = \sum_{j=m+G, \dots, m+2G-1} \{ A1(2, j)^2 + B1(2, j)^2 \}$$

埋め込みデータが値 1 の場合、 $y = (E1d \cdot E2u) / (E1u \cdot E2d)$

埋め込みデータが値 2 の場合、 $y = (E1u \cdot E2d) / (E1d \cdot E2u)$

【 0 0 5 0 】

40

上記〔数式 6〕において、 m は変更対象周波数帯の下限の成分の番号、 $m + 2G$ は変更対象周波数帯の上限の成分の番号である。例えば、変更対象周波数帯として、1 . 7 k H z ~ 3 . 4 k H z を設定する場合、 $m = 160$ 、 $m + 2G = 320$ となる。したがって、1 つの周波数領域の幅 $G = 80$ である。

【 0 0 5 1 】

そして、さらに強度割合 y の値に応じて、周波数成分変更手段 3 0 は、以下の〔数式 7〕に従った処理を実行することにより、係数 α 、 β を補正して係数 α' 、 β' を得る。

【 0 0 5 2 】

〔数式 7〕

$$0.01 \quad y < 1.0 \text{ の場合、 } \alpha' = \alpha \cdot y^{-1/2}、\beta' = \beta \cdot y^{1/2}$$

50

$y < 0.01$ の場合、 $\alpha' = 10.0 \cdot \alpha$ 、 $\beta' = 0.1 \cdot \beta$

【0053】

なお、 $y = 1.0$ の場合は、補正を行わない。さらに、周波数成分変更手段30は、埋め込むべき情報が“値1”である場合、以下の〔数式8〕に従った処理を実行することにより、周波数成分の状態を“状態1”、すなわち、図5(c)(e)に示したような状態に変更する。

【0054】

〔数式8〕

$j = m \sim m + G - 1$ の各成分に対して

$$E(1, j) = \{A_l(1, j)^2 + B_l(1, j)^2 + A_r(1, j)^2 + B_r(1, j)^2\}^{1/2} \quad 10$$

$$A_l'(1, j) = A_l(1, j) \cdot E(1, j) \cdot \alpha / \{A_l(1, j)^2 + B_l(1, j)^2\}^{1/2}$$

$$B_l'(1, j) = B_l(1, j) \cdot E(1, j) \cdot \alpha / \{A_l(1, j)^2 + B_l(1, j)^2\}^{1/2}$$

$$A_r'(1, j) = A_r(1, j) \cdot E(1, j) \cdot \alpha / \{A_r(1, j)^2 + B_r(1, j)^2\}^{1/2}$$

$$B_r'(1, j) = B_r(1, j) \cdot E(1, j) \cdot \alpha / \{A_r(1, j)^2 + B_r(1, j)^2\}^{1/2}$$

$$E(2, j) = \{A_l(2, j)^2 + B_l(2, j)^2 + A_r(2, j)^2 + B_r(2, j)^2\}^{1/2} \quad 20$$

$$A_l'(2, j) = A_l(2, j) \cdot E(2, j) \cdot \beta / \{A_l(2, j)^2 + B_l(2, j)^2\}^{1/2}$$

$$B_l'(2, j) = B_l(2, j) \cdot E(2, j) \cdot \beta / \{A_l(2, j)^2 + B_l(2, j)^2\}^{1/2}$$

$$A_r'(2, j) = A_r(2, j) \cdot E(2, j) \cdot \beta / \{A_r(2, j)^2 + B_r(2, j)^2\}^{1/2}$$

$$B_r'(2, j) = B_r(2, j) \cdot E(2, j) \cdot \beta / \{A_r(2, j)^2 + B_r(2, j)^2\}^{1/2}$$

$j = m + G \sim m + 2G - 1$ の各成分に対して 30

$$E(1, j) = \{A_l(1, j)^2 + B_l(1, j)^2 + A_r(1, j)^2 + B_r(1, j)^2\}^{1/2}$$

$$A_l'(1, j) = A_l(1, j) \cdot E(1, j) \cdot \beta / \{A_l(1, j)^2 + B_l(1, j)^2\}^{1/2}$$

$$B_l'(1, j) = B_l(1, j) \cdot E(1, j) \cdot \beta / \{A_l(1, j)^2 + B_l(1, j)^2\}^{1/2}$$

$$A_r'(1, j) = A_r(1, j) \cdot E(1, j) \cdot \beta / \{A_r(1, j)^2 + B_r(1, j)^2\}^{1/2}$$

$$B_r'(1, j) = B_r(1, j) \cdot E(1, j) \cdot \beta / \{A_r(1, j)^2 + B_r(1, j)^2\}^{1/2} \quad 40$$

$$E(2, j) = \{A_l(2, j)^2 + B_l(2, j)^2 + A_r(2, j)^2 + B_r(2, j)^2\}^{1/2}$$

$$A_l'(2, j) = A_l(2, j) \cdot E(2, j) \cdot \alpha / \{A_l(2, j)^2 + B_l(2, j)^2\}^{1/2}$$

$$B_l'(2, j) = B_l(2, j) \cdot E(2, j) \cdot \alpha / \{A_l(2, j)^2 + B_l(2, j)^2\}^{1/2}$$

$$A_r'(2, j) = A_r(2, j) \cdot E(2, j) \cdot \alpha / \{A_r(2, j)^2 + B_r(2, j)^2\}^{1/2}$$

$$B_r'(2, j) = B_r(2, j) \cdot E(2, j) \cdot \alpha / \{A_r(2, j)^2 + B_r(2, j)^2\}^{1/2} \quad 50$$

【 0 0 5 5 】

埋め込むべき情報が“値 2”である場合は、以下の〔数式 9〕に従った処理を実行することにより、周波数成分の状態を“状態 2”、すなわち、図 5 (d) (f) に示したような状態に変更する。

【 0 0 5 6 】

〔数式 9〕

$j = m \sim m + G - 1$ の各成分に対して

$$E(1, j) = \{ A_l(1, j)^2 + B_l(1, j)^2 + A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2} \quad 10$$

$$A_l'(1, j) = A_l(1, j) \cdot E(1, j) \cdot \beta / \{ A_l(1, j)^2 + B_l(1, j)^2 \}^{1/2}$$

$$B_l'(1, j) = B_l(1, j) \cdot E(1, j) \cdot \beta / \{ A_l(1, j)^2 + B_l(1, j)^2 \}^{1/2}$$

$$A_r'(1, j) = A_r(1, j) \cdot E(1, j) \cdot \beta / \{ A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2}$$

$$B_r'(1, j) = B_r(1, j) \cdot E(1, j) \cdot \beta / \{ A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2}$$

$$E(2, j) = \{ A_l(2, j)^2 + B_l(2, j)^2 + A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2} \quad 20$$

$$A_l'(2, j) = A_l(2, j) \cdot E(2, j) \cdot \alpha / \{ A_l(2, j)^2 + B_l(2, j)^2 \}^{1/2}$$

$$B_l'(2, j) = B_l(2, j) \cdot E(2, j) \cdot \alpha / \{ A_l(2, j)^2 + B_l(2, j)^2 \}^{1/2}$$

$$A_r'(2, j) = A_r(2, j) \cdot E(2, j) \cdot \alpha / \{ A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2}$$

$$B_r'(2, j) = B_r(2, j) \cdot E(2, j) \cdot \alpha / \{ A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2}$$

$j = m + G \sim m + 2G - 1$ の各成分に対して

$$E(1, j) = \{ A_l(1, j)^2 + B_l(1, j)^2 + A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2} \quad 30$$

$$A_l'(1, j) = A_l(1, j) \cdot E(1, j) \cdot \alpha / \{ A_l(1, j)^2 + B_l(1, j)^2 \}^{1/2}$$

$$B_l'(1, j) = B_l(1, j) \cdot E(1, j) \cdot \alpha / \{ A_l(1, j)^2 + B_l(1, j)^2 \}^{1/2}$$

$$A_r'(1, j) = A_r(1, j) \cdot E(1, j) \cdot \alpha / \{ A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2}$$

$$B_r'(1, j) = B_r(1, j) \cdot E(1, j) \cdot \alpha / \{ A_r(1, j)^2 + B_r(1, j)^2 \}^{1/2}$$

$$E(2, j) = \{ A_l(2, j)^2 + B_l(2, j)^2 + A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2} \quad 40$$

$$A_l'(2, j) = A_l(2, j) \cdot E(2, j) \cdot \beta / \{ A_l(2, j)^2 + B_l(2, j)^2 \}^{1/2}$$

$$B_l'(2, j) = B_l(2, j) \cdot E(2, j) \cdot \beta / \{ A_l(2, j)^2 + B_l(2, j)^2 \}^{1/2}$$

$$A_r'(2, j) = A_r(2, j) \cdot E(2, j) \cdot \beta / \{ A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2}$$

$$B_r'(2, j) = B_r(2, j) \cdot E(2, j) \cdot \beta / \{ A_r(2, j)^2 + B_r(2, j)^2 \}^{1/2}$$

【 0 0 5 7 】

なお、上記〔数式 7〕に従った処理を実行して係数 α' 、 β' を得た場合は、上記〔数 50

式 8〕〔数式 9〕において、係数 α 、 β に代えて係数 α' 、 β' を用いる。

【 0 0 5 8 】

周波数逆変換手段 40 は、上記のようにして、周波数成分の状態が変更されたフレームスペクトルを周波数逆変換して改変音響フレームを得る処理を行う。この周波数逆変換は、当然のことながら、周波数変換手段 20 が実行した手法に対応していることが必要となる。本実施形態では、周波数変換手段 20 において、フーリエ変換を施しているため、周波数逆変換手段 40 は、フーリエ逆変換を実行することになる。

【 0 0 5 9 】

具体的には、A タイプの音響フレームに対しては、周波数逆変換手段 40 は、上記〔数式 8〕〔数式 9〕のいずれかにより得られたスペクトルの左チャンネルの実部 $A l'(1, j)$ 等、虚部 $B l'(1, j)$ 等、右チャンネルの実部 $A r'(1, j)$ 等、虚部 $B r'(1, j)$ 等を用いて、以下の〔数式 10〕に従った処理を行い、 $X l'(i)$ 、 $X r'(i)$ を算出する。なお、上記〔数式 8〕〔数式 9〕において改変されていない周波数成分については、 $A l'(1, j)$ 等として、元の周波数成分である $A l(1, j)$ 等を用いる。周波数逆変換を計算するにあたり、 $A l'(1, j)$ および $B l'(1, j)$ に対しては〔数式 5〕における $Z l(1)$ を、 $A r'(1, j)$ および $B r'(1, j)$ に対しては〔数式 5〕における $Z r(1)$ を除することにより、同時に振幅逆変換を施す必要がある。

【 0 0 6 0 】

〔数式 10〕

$$X l'(i) = 1 / N \cdot \{ \sum_j A l'(1, j) \cdot \cos(2 \pi i j / N) / Z l(1) - \sum_j B l'(1, j) \cdot \sin(2 \pi i j / N) / Z l(1) \} + X l p(i + N / 2)$$

$$X r'(i) = 1 / N \cdot \{ \sum_j A r'(1, j) \cdot \cos(2 \pi i j / N) / Z r(1) - \sum_j B r'(1, j) \cdot \sin(2 \pi i j / N) / Z r(1) \} + X r p(i + N / 2)$$

【 0 0 6 1 】

上記〔数式 10〕においては、式が繁雑になるのを防ぐため、 $\sum_{j=0, \dots, N-1}$ を $\sum_j$ と示している。上記〔数式 10〕における第 1 式の “ $+ X l p(i + N / 2)$ ”、第 2 式の “ $+ X r p(i + N / 2)$ ” の項は、直前に改変された改変音響フレームのデータ $X l p(i)$ 、 $X r p(i)$ が存在する場合に、時間軸上 $N / 2$ サンプル分重複することを考慮して加算するためのものである。上記〔数式 10〕により A タイプの改変音響フレームの左チャンネルの各サンプル $X l'(i)$ 、右チャンネルの各サンプル $X r'(i)$ が得られることになる。

【 0 0 6 2 】

B タイプの音響フレームに対しては、周波数逆変換手段 40 は、上記〔数式 8〕〔数式 9〕のいずれかにより得られたスペクトルの左チャンネルの実部 $A l'(2, j)$ 、虚部 $B l'(2, j)$ 、右チャンネルの実部 $A r'(2, j)$ 、虚部 $B r'(2, j)$ を用いて、以下の〔数式 11〕に従った処理を行い、 $X l'(i)$ 、 $X r'(i)$ を算出する。なお、上記〔数式 8〕〔数式 9〕において改変されていない周波数成分については、以下の〔数式 11〕においては $A l'(2, j)$ 、 $B l'(2, j)$ 、 $A r'(2, j)$ 、 $B r'(2, j)$ として、元の値である $A l(2, j)$ 、 $B l(2, j)$ 、 $A r(2, j)$ 、 $B r(2, j)$ を用いる。周波数逆変換を計算するにあたり、 $A l'(2, j)$ および $B l'(2, j)$ に対しては〔数式 5〕における $Z l(2)$ を、 $A r'(2, j)$ および $B r'(2, j)$ に対しては〔数式 5〕における $Z r(2)$ を除することにより、同時に振幅逆変換を施す必要がある。

【 0 0 6 3 】

〔数式 11〕

$$X l'(i + N / 2) = 1 / N \cdot \{ \sum_j A l'(2, j) \cdot \cos(2 \pi i j / N) / Z l(2) - \sum_j B l'(2, j) \cdot \sin(2 \pi i j / N) / Z l(2) \} + X l p(i + N / 2)$$

10

20

30

40

50

$i + N)$

$$Xr'(i + N/2) = 1/N \cdot \{ \sum_j Ar'(2, j) \cdot \cos(2\pi i j / N) / Zr(2) - \sum_j Br'(2, j) \cdot \sin(2\pi i j / N) / Zr(2) \} + Xrp(i + N)$$

【0064】

上記〔数式11〕によりBタイプの改変音響フレームの左チャンネルの各サンプル $Xl'(i)$ 、右チャンネルの各サンプル $Xr'(i)$ 、が得られることになる。

【0065】

改変音響フレーム出力手段50は、周波数逆変換手段40の処理により得られたAタイプの改変音響フレーム、Bタイプの改変音響フレームを順次出力ファイルに出力する。

10

【0066】

次に、図2に示した音響信号に対する情報の埋め込み装置の処理の全体的な流れについて説明する。埋め込み装置を起動すると、まず、変換テーブル作成手段80が、 Nw ビットの基準符号と Nh ビットのハミング符号が対応付けられた符号変換テーブルを作成する。変換テーブル作成手段80による符号変換テーブル作成のフローチャートを図6に示す。

【0067】

変換テーブル作成手段80は、最初に初期化処理を行う(S601)。具体的には、7ビット符号“0”に12ビット符号“1”を対応付けて符号変換テーブルの $i (= 0)$ 番目に登録するとともに、7ビット符号 KF の初期値を1、12ビット符号 HF の初期値を2とする。続いて、 $i = 0$ に初期設定する(S602)。次に、12ビット符号 HF と符号変換テーブルの i 番目に既に登録されている他の12ビット符号とのハミング距離を算出する(S603)。

20

【0068】

算出したハミング距離が4未満の場合、 HF の値を1増加して更新した後(S604)、S602に戻って、次の12ビット符号 HF についての処理を行う。一方、算出したハミング距離が4以上の場合、 i の値を1増加して更新する(S605)。そして、 i が $KF - 1$ 未満である場合には、S603に戻って、符号変換テーブル内の i 番目に登録されている12ビット符号とのハミング距離を算出する。S605の後、 i が $KF - 1$ 以上となった場合には、12ビット符号 HF を符号変換テーブルの7ビット符号 KF の位置に登録するとともに、 KF 、 HF の値をそれぞれ1ずつ増加して更新する(S606)。そして、 KF が128未満である場合は、S602に戻って、次の7ビット符号 KF についての処理を行う。 KF が128以上である場合は、全ての7ビット符号 KF に対応する12ビット符号 HF が登録されたことになるので、符号変換テーブル作成処理を終了する。

30

【0069】

このようにして作成された符号変換テーブルを図7に示す。図7に示すように、符号変換テーブルには、7ビット符号が取り得る各値0～127について、12ビット符号が対応付けて登録されている。12ビット符号の2進表記を見るとわかるように、全ての12ビット符号は互いにハミング距離が4以上となっている。なお、図7の例では、説明の便宜上、7ビット符号の10進表記と、12ビット符号の10進表記及び2進表記を示しているが、現実には、7ビット符号のビット列と12ビット符号のビット列が対応付けて登録されている。

40

【0070】

続いて、符号変換テーブル作成処理後の埋め込み装置の処理を、図8のフローチャートに従って説明する。図8は、付加情報1ワード分に対応した処理を示している。本発明では、1ワード分を設定された所定回数 Na だけ繰り返し実行した後、次の1ワードを読み込む処理を実行する。例えば、 $Na = 3$ と設定されている場合、3ワード分同じビット列が繰り返し埋め込まれる。なお、1ワードのビット数 Nw としては、任意のビット数に設定することができるが、上述のように、本実施形態では、ASCIIコードの実質7ビットに設定されている。

50

【 0 0 7 1 】

図 8 においては、まず、ビット配列作成手段 7 0 は、付加情報記憶部 6 2 から抽出した付加情報の各ワードについて、変換テーブル作成手段 8 0 を参照し、対応するビット配列を作成する (S 1 0 1)。具体的には、まず、付加情報記憶部 6 2 から 1 ワード (7 ビット) 単位で抽出し、図 7 に示した符号変換テーブルを参照し、対応するハミング符号で構成される 1 2 ビットのビット配列を抽出する。

【 0 0 7 2 】

そして、この 1 2 ビットが、音響信号に対する情報の埋め込み装置として用いられるコンピュータ内のレジスタに読み込まれることになる。このように、付加情報記憶部 6 2 においては、1 ワードは 7 ビットであるが、埋め込み処理時は、この 1 2 ビットの配列で、付加情報内の 1 ワード分の処理を行う。

【 0 0 7 3 】

次に、周波数成分変更手段 3 0 が、レジスタに保持された $Nh (= 12)$ ビットから 1 ビットを読み込む処理を行う (S 1 0 2)。続いて、音響フレーム読込手段 1 0 が、音響信号記憶部 6 1 に記憶されたステレオ音響信号の左右の各チャンネルから、それぞれ所定数のサンプルを A タイプの 1 音響フレームとして読み込み、周波数変換手段 2 0 が周波数変換を行って、得られたフレームスペクトルに対して振幅変換を行う (S 1 0 3)。具体的には、まず、読み込んだ音響フレームに対して、周波数変換を行って、その音響フレームのスペクトルであるフレームスペクトルを得る。すなわち、各音響フレームについて、窓関数 $W(1, i)$ を用いて、上記〔数式 2〕に従った処理を行う。そして、上記〔数式 5〕に従った処理を実行して $Zl(1)$ 、 $Zr(1)$ を算出し振幅変換を行う。同様に、音響フレーム読込手段 1 0 が、音響信号記憶部 6 1 に記憶されたステレオ音響信号の左右の各チャンネルから、それぞれ所定数のサンプルを B タイプの 1 音響フレームとして読み込み、周波数変換手段 2 0 が周波数変換を行って、得られたフレームスペクトルに対して振幅変換を行う (S 1 0 4)。具体的には、まず、読み込んだ音響フレームに対して、周波数変換を行って、その音響フレームのスペクトルであるフレームスペクトルを得る。すなわち、各音響フレームについて、窓関数 $W(2, i)$ を用いて、上記〔数式 3〕に従った処理を行う。そして、上記〔数式 5〕に従った処理を実行して $Zl(2)$ 、 $Zr(2)$ を算出し振幅変換を行う。

【 0 0 7 4 】

続いて、周波数成分変更手段 3 0 が、読み込んだビット値に応じて A タイプ音響フレームおよび B タイプ音響フレームの周波数成分の状態を変更するにあたり、〔数式 6〕〔数式 7〕に従った処理を実行して変換割合 α 、 β を決定する処理を行う (S 1 0 5)。この決定された変換割合を用いて、周波数成分変更手段 3 0 は、ビット配列作成手段 7 0 から受け取った値 1、値 2 に応じて上記〔数式 8〕〔数式 9〕に従った処理を実行し、変更対象周波数帯の成分の状態を“状態 1”、“状態 2”に応じた状態のいずれかに変更する (S 1 0 6)。

【 0 0 7 5 】

次に、周波数逆変換手段 4 0 が、上記 S 1 0 6 の処理により A タイプの音響フレームに対応する各スペクトル集合の強度が変更されたスペクトルを振幅逆変換、周波数逆変換して改変音響フレームを得る処理を行う (S 1 0 7)。この振幅逆変換は〔数式 5〕で算出した $Zl(1)$ 、 $Zr(1)$ の逆数をスペクトルに乗じることにより行なう。この周波数逆変換は、当然のことながら、周波数変換手段 2 0 が S 1 0 3 において実行した手法に対応していることが必要となる。本実施形態では、周波数変換手段 2 0 において、フーリエ逆変換を施しているため、周波数逆変換手段 4 0 は、フーリエ逆変換を実行することになる。具体的には、上記〔数式 8〕〔数式 9〕のいずれかにより得られたスペクトルの左チャンネルの実部 $A1'(1, j)$ 等、虚部 $B1'(1, j)$ 等、右チャンネルの実部 $Ar'(1, j)$ 等、虚部 $Br'(1, j)$ 等を用いて、上記〔数式 10〕に従った処理を行い、 $Xl'(i)$ 、 $Xr'(i)$ を算出する。改変音響フレーム出力手段 5 0 は、得られた改変音響フレームを順次出力ファイルに出力する。

【 0 0 7 6 】

同様に、周波数逆変換手段 4 0 が、上記 S 1 0 6 の処理により B タイプの音響フレームに対応する各スペクトル集合の強度が変更されたスペクトルを振幅逆変換、周波数逆変換して改変音響フレームを得る処理を行う (S 1 0 8)。具体的には、振幅逆変換は〔数式 5〕で算出した $Z_l(2)$ 、 $Z_r(2)$ の逆数をスペクトルに乗じることにより行ない、上記〔数式 8〕〔数式 9〕のいずれかにより得られたスペクトルの左チャンネルの実部 $A_{l'}(2, j)$ 等、虚部 $B_{l'}(2, j)$ 等、右チャンネルの実部 $A_{r'}(2, j)$ 等、虚部 $B_{r'}(2, j)$ 等を用いて、上記〔数式 1 1〕に従った処理を行い、 $X_{l'}(i)$ 、 $X_{r'}(i)$ を算出する。

【 0 0 7 7 】

改変音響フレーム出力手段 5 0 は、得られた改変音響フレームを順次出力ファイルに出力する。こうして各チャンネルについて、A タイプおよび B タイプの 2 つの音響フレームに対する処理を終えたら、周波数成分変更手段 3 0 がビット配列中の次の 1 ビットを読み込む (S 1 0 2)。以上のような処理を音響信号の両チャンネルの全サンプルに渡って実行していく。すなわち、所定数のサンプルを音響フレームとして読み込み、音響信号から読み込むべき音響フレームがなくなったら (S 1 0 3、S 1 0 4)、処理を終了する。なお、S 1 0 1 において読み込んだ 1 ワード分のビット配列 ($N_h = 12$ ビット) の各ビットに対応する処理を終えた場合、設定された所定回数 N_a に達していないときは、先頭のビットに戻って S 1 0 3 ~ S 1 0 8 の処理を繰り返し実行する。設定された所定回数 N_a に達したときは、S 1 0 2 から S 1 0 1 に戻り、付加情報の次のワードを読み込んでビット配列を作成する処理を行うことになる。付加情報の全ワードに対して処理が終了した場合は、付加情報の先頭ワードに戻って処理を行う。この結果、全ての音響フレームに対して処理を行った全ての改変音響フレームが出力ファイルに記録されて、改変音響信号として得られる。得られた改変音響信号は、記憶手段 6 0 内の改変音響信号記憶部 6 3 に出力され、記憶される。

【 0 0 7 8 】

なお、本実施形態では、付加情報を 1 ワード 7 ビットとし、符号変換テーブルにより 12 ビットのビット配列に変換して、付加情報 1 ワード分の処理をする場合について説明したが、本発明では、抽出側と取り決めがある限り、付加情報の 1 ワードを他のビット数単位で記録することが可能である。

【 0 0 7 9 】

上記のようにして得られた改変音響信号のうち、付加情報が埋め込まれている部分については、変更対象周波数帯の成分は、上記状態 1、状態 2 の 2 通りの分布しかないのである。しかし、変更対象周波数帯の成分以外については、元の音響信号のままであるので、制作者の設定に基づいた種々な分布になる。また、上記〔数式 8〕〔数式 9〕の処理からも明らかなように、変更対象周波数帯の成分は、左チャンネルと右チャンネルにおいて同等の割合で変化させられている。したがって、両スピーカと等距離の位置においても、変更対象周波数帯の成分は、相殺されることなく増幅させる関係となり、情報を抽出し易くなる。

【 0 0 8 0 】

以上、埋め込み装置について説明したが、ここで、本実施形態である振幅変換を用いた埋め込みおよび抽出の手法について図 9 から図 1 2 を用いて説明する。図 9 および図 1 0 は、比較のために、振幅変換を用いない場合の埋め込み処理および抽出処理を示し、図 1 1 および図 1 2 は本実施形態における振幅変換を用いた埋め込み処理および抽出処理を示す。これらの説明図においては、音響フレーム奇数および偶数 3 対分について、3 ビットの情報を埋め込むことを想定した音響信号の信号波形を示しており、説明の便宜上、音響フレームが互いに重複しない状態で示している。各図において、左端の波形は時間次元の通常の音響信号波形を示しており、横方向は時間軸であり、右に向かうに従って時間が進行し、縦方向は強度 (振幅) である。中央または右端の波形は周波数変換後の所定の周波数範囲のスペクトル成分の合算値を信号波形で表現しており、仮想的にバンドパスフィル

10

20

30

40

50

ターを通した後の信号波形になっている。縦方向は同様に強度であり、本来は左端に比べ顕著に低くなるが、ここでは説明の都合上、左端と同程度に拡大して表現している。また、2種類の波形のレベルも本来は差があるが、ここでは同一レベルで表現している。実際には、周波数変換後のデータとして所定の周波数範囲外の高周波数成分や低周波数成分も存在するが、これらについては改変を加えないという前提で図示しておらず、図5に従った埋め込みを行なう2種の周波数帯域成分に限定して2種の信号波形で示している。

【0081】

図9(a)は原音響信号を示しており、最初の1対の音響フレームは全体に渡って強度がほぼ一定、中央の1対の音響フレームは奇数フレームの強度が小さく偶数フレームの強度が大きいものであり、最後の1対の音響フレームは奇数フレームの強度が大きく偶数フレームの強度が小さいものである。このような音響信号に対して、周波数変換を行った結果が図9(b)で、図5で示されように埋め込み対象成分は上下2分割して成分変更を施す都合上、図9(b)では(b-1)および(b-2)の2種の波形に分離して表現した。周波数変換後の波形形状は周波数ごとに異なり、図9(a)と必ずしも相似形にはならないが、ここでは周波数変換前と同様になると仮定して説明する。図9(b-1)および(b-2)に対して、[0, 0, 0]という3ビットのデータを図5で説明した方法に基づいて埋め込むことにより(図5ではビット0は符号1に相当する)、図9(c-1)および(c-2)が得られる。最初の1対は図5のようなパターンが表現できているが、残り2対の上一方は、原段階に対して逆方向に大小関係を構築することが要求され、品質維持の都合上、適切なパターンが構築できていないことがわかる。このような改変を加えた結果に対して、周波数逆変換を行った結果が図9(d)である。一般に音響信号波形の外形(エンベロープという)はエネルギー分布が大きい低周波成分で決定され、これらは本願では所定の周波数範囲外の信号成分に含まれ、図9(c)の段階では改変が加えられていないため、周波数逆変換後の図9(d)は周波数変換前の図9(a)と類似した形状になる。

【0082】

このような埋め込みを行なった結果に対して、抽出処理を行う様子を図10に示す。図10(a)および図10(b)は各々図9(d)および図9(c)に対応する。図10(b)において、ビット判定を行なうにあたり、最初の1対以外は図5で想定されているパターンが形成されていないため、正しいビットを判定することが難しくなる。

【0083】

次に、本実施形態における振幅変換を導入した手法について図11および図12を用いて説明する。図11(a)は図9(a)と同じ原音響信号を示しており、同様に周波数変換を行った図11(b)も図9(b)と同じである。ここで、6つのフレーム単位に振幅変換を行った結果が図11(c)である。図11の場合、図11(a)で各フレーム内の振幅が平坦であるため、図11(c)では全体が平坦になっているが、実際には各フレーム内は微細に変動しており、そのフレーム内変動分は図11(c)の段階でも踏襲されるため、本図のように完全に平坦になることは通常ない。(実際は図11(c)は周波数次元のデータになるため、周波数次元のデータも平坦にならず変動した形態になる。)また、変換倍率はフレームごとに設定し、上下2つの周波数成分に対しては同一の倍率で変換を行うため、通常は上下で顕著な差が生じる(ただし、時間軸方向は比較的揃う)。これに対して、[0, 0, 0]という3ビットのデータを埋め込むと、図11(d-1)および(d-2)が得られる。原段階である図11(c)が平坦な波形であるため、全てのフレームにおいて図5のような理想的なパターンが容易に構築できることがわかる。続いて、前記フレームごとに設定した倍率の逆数を乗じて振幅逆変換を行った結果が図11(e)である。この段階では図9(c)と類似したパターンが生じることが多いが、本形状はいかなるものでも構わない。最後に周波数逆変換を行うと、図11(f)が得られ、同様に原信号波形の図11(a)と類似した形状になる。

【0084】

このような埋め込みを行なった結果に対して、抽出処理を行う様子を図12に示す。図

10

20

30

40

50

12 (a)、(b)および(c)は各々図11(f)、(e)および(d)に対応する。図12(b)の周波数変換後の波形形状は図11(b)とは基本的に異なるが、算出される振幅変換倍率は類似した値になり、ほぼ同様な倍率で振幅変換が行われ図12(c)が得られる。図12(c)の段階で、ビット判定を行なうと、全てのフレームにおいて図5のような理想的なパターンが形成されているため、正しいビットを判定することが可能になる。

【0085】

以上、本実施形態では、図11(c)に示したように、埋め込み対象の信号成分が時間軸方向に平坦になるように変換を行って埋め込むようにしているため、高周波側と低周波側の成分強度の大小関係が完全に逆になるような不自然な変更を行う確率が低くなり、品質を維持しながら、抽出側における抽出精度を高めることが可能となる。

10

【0086】

(3.1.抽出装置の構成)

次に、本発明に係る音響信号からの情報の抽出装置について説明する。図13は、本発明に係る音響信号からの情報の抽出装置の一実施形態を示す構成図である。図13において、100は音響信号入力手段、110は基準フレーム獲得手段、120は位相変更フレーム設定手段、130は周波数変換手段、140は符号判定パラメータ算出手段、150は符号出力手段、160は付加情報抽出手段、170は音響フレーム保持手段、180は変換テーブル作成手段である。

【0087】

20

音響信号入力手段100は、流れている音声をデジタル音響信号として取得し、入力する機能を有している。現実には、マイクロフォンおよびA/D変換器により実現される。マイクロフォンとしては、変更対象周波数帯の成分が検出可能なものであれば、モノラル無指向性のものであっても、ステレオ指向性のものであっても使用可能である。ステレオ指向性のものであっても一方のチャンネルだけ利用すれば良い。また、図2に示した装置で情報の埋め込みを行った場合には、ステレオで再生されるが、左右どちらのスピーカの音を入力しても、双方のスピーカの音を混合入力しても良く、マイクロフォンの設置位置に制約がない。もちろん、付加情報が埋め込まれた音響信号がモノラル再生される場合、あるいは付加情報が埋め込まれた音響信号自体がモノラルである場合は、再生される単一のスピーカにマイクロフォンを向ければ良い。このマイクロフォンは特別精度の高いものでなく、一般的な精度のマイクロフォンを用いても情報の抽出が可能となる。基準フレーム獲得手段110は、入力されたデジタルのモノラル音響信号(あるいはステレオ音響信号の1チャンネル)から所定数のサンプルで構成される音響フレームを基準フレームとして読み込む機能を有している。基準フレームとしては、埋め込み時と同様にAタイプ、Bタイプのもので設定される。位相変更フレーム設定手段120は、Aタイプ、Bタイプそれぞれの基準フレームと所定サンプルずつ移動させることにより位相を変更した音響フレームを位相変更フレームとして設定する機能を有している。

30

【0088】

周波数変換手段130は、基本的には、図2に示した周波数変換手段20とほぼ同様の機能を有している。ただし、音を取り込むタイミングが、音響信号の先頭からとはならないため、正しい位相を特定するために、複数の位相で振幅変換、周波数変換を行う点、元の音響信号がステレオであっても、1つのチャンネルで行う点が異なっている。

40

【0089】

周波数変換手段130は、Aタイプの音響フレームに対してフーリエ変換を行う場合は、信号 $X(i - N/2 + pN/6)$ ($i = 0, \dots, N - 1$) に対して、窓関数 $W(1, i)$ を用いて、以下の〔数式12〕に従った処理を行い、変換データの実部 $A(1, j, p)$ 、虚部 $B(1, j, p)$ を得る機能を有している。 p は位相番号であり、0~5の整数値をとる。

【0090】

〔数式12〕

50

$$A(1, j, p) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X(i - N/2 + p \cdot N/6) \cdot \cos(2\pi i j / N)$$

$$B(1, j, p) = \sum_{i=0, \dots, N-1} W(1, i) \cdot X(i - N/2 + p \cdot N/6) \cdot \sin(2\pi i j / N)$$

【0091】

周波数変換手段130は、Bタイプの音響フレームに対してフーリエ変換を行う場合は、信号 $X(i + p \cdot N/6)$ ($i = 0, \dots, N-1$)に対して、窓関数 $W(2, i)$ を用いて、以下の〔数式13〕に従った処理を行い、変換データの実部 $A(2, j, p)$ 、虚部 $B(2, j, p)$ を得る機能を有している。

【0092】

10

〔数式13〕

$$A(2, j, p) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X(i + p \cdot N/6) \cdot \cos(2\pi i j / N)$$

$$B(2, j, p) = \sum_{i=0, \dots, N-1} W(2, i) \cdot X(i + p \cdot N/6) \cdot \sin(2\pi i j / N)$$

【0093】

周波数変換手段130は、周波数変換手段20と同様に振幅変換を行う。振幅変換を行うにあたり、まず変換倍率の算出を行う。変換倍率の算出は、所定周波数範囲のスペクトル集合の実効強度値の平均である平均実効値で、設定値を除算することにより行う。設定値は、適宜定めておくことができるが、埋め込み時における振幅変換の場合と同一値としておくことが必要となる。したがって、本実施形態では、設定値 Z_0 は160にする必要がある。具体的には、Aタイプの音響フレーム、Bタイプの音響フレームについての変換倍率 $Z(1, p)$ 、 $Z(2, p)$ 、これらの直前のAタイプの音響フレーム、Bタイプの音響フレームについての変換倍率 $Z_{-1}(1, p)$ 、 $Z_{-1}(2, p)$ は、以下の〔数式14〕に従った処理により算出される。振幅変換が実行されることにより、全ての音響フレームは、元の状態における信号強度に関わらず、各音響フレームの平均実効値が設定値となるように振幅変換された状態で、情報の抽出が行われることになる。なお、 p は位相番号であり、0～5の整数値をとる。

20

【0094】

〔数式14〕

30

$$Z(1, p) = Z_0 / \left[\sum_{j=m, \dots, m+2G-1} \{ A(1, j, p)^2 + B(1, j, p)^2 \} \right]^{1/2}$$

$$Z(2, p) = Z_0 / \left[\sum_{j=m, \dots, m+2G-1} \{ A(2, j, p)^2 + B(2, j, p)^2 \} \right]^{1/2}$$

$$Z_{-1}(1, p) = Z_0 / \left[\sum_{j=m, \dots, m+2G-1} \{ A_{-1}(1, j, p)^2 + B_{-1}(1, j, p)^2 \} \right]^{1/2}$$

$$Z_{-1}(2, p) = Z_0 / \left[\sum_{j=m, \dots, m+2G-1} \{ A_{-1}(2, j, p)^2 + B_{-1}(2, j, p)^2 \} \right]^{1/2}$$

【0095】

40

$j = m, \dots, m + 2G - 1$ (周波数 $F_1, \dots, F_2$ に相当)および $p = 0, \dots, 5$ の範囲で、 $A(1, j, p)$ および $B(1, j, p)$ の各々の要素に対して $Z(1, p)$ を乗じ、 $A(2, j, p)$ および $B(2, j, p)$ の各々の要素に対して $Z(2, p)$ を乗じ、 $A_{-1}(1, j, p)$ および $B_{-1}(1, j, p)$ の各々の要素に対して $Z_{-1}(1, p)$ を乗じ、 $A_{-1}(2, j, p)$ および $B_{-1}(2, j, p)$ の各々の要素に対して $Z_{-1}(2, p)$ を乗じることにより、振幅変換を行う。以下説明において、 $A(1, j, p)$ 、 $B(1, j, p)$ 、 $A(2, j, p)$ 、 $B(2, j, p)$ 、 $A_{-1}(1, j, p)$ 、 $B_{-1}(1, j, p)$ 、 $A_{-1}(2, j, p)$ 、 $B_{-1}(2, j, p)$ はこれらの振幅変換を行った値とする。

【0096】

符号判定パラメータ算出手段140は、生成されたスペクトルから所定の周波数範囲に

50

相当するスペクトル集合を抽出し、各スペクトル集合の強度値を算出するとともに、この強度値を利用して符号判定パラメータを算出し、この符号判定パラメータの大小関係に基づいて、所定の状態であると判断する機能を有している。上述のように、本実施形態では、Aタイプの音響フレームとBタイプの音響フレームが $N/2$ サンプルずつ重複して設定されているため、ある音響フレームについて、強度値、符号判定パラメータを算出する際には、直前の音響フレームによる残響成分を考慮する必要がある。ところが、残響成分は計算により算出されるものであるもので、必ずしも正確なものが算出されるとは限らず、算出された残響成分を除去することにより、反って抽出精度が低下する場合がある。そこで、本実施形態では、残響成分を除去しない場合の強度値 E_1 、 E_2 、 E_3 、 E_4 、および符号判定パラメータ C と、除去した場合の強度値 E_1' 、 E_2' 、 E_3' 、 E_4' 、および補正符号判定パラメータ C' を算出し、これらを利用して最適であると思われる状態を判断する。

10

20

30

40

50

【0097】

Aタイプの音響フレームの各成分 $A(1, j, p)$ 、 $B(1, j, p)$ に対して1フレーム分前のBタイプの音響フレームに対応する各成分を $A_{-1}(2, j, p)$ 、 $B_{-1}(2, j, p)$ とすると、Bタイプの音響フレームの各成分 $A(2, j, p)$ 、 $B(2, j, p)$ に対応して、1フレーム分前の音響フレームは以下補正を施す前の前記Aタイプの音響フレームとなる。符号判定パラメータ算出手段140は、各スペクトル集合の強度値の算出の前に、まず、各スペクトル集合の基本強度値 $E(1, j, p)$ 、 $E(2, j, p)$ 、および残響成分を除去した基本強度値 $E'(1, j, p)$ 、 $E'(2, j, p)$ を、以下の〔数式15〕に従った処理により算出する。

【0098】

〔数式15〕

$$\begin{aligned} E(1, j, p) &= A(1, j, p)^2 + B(1, j, p)^2 \\ E(2, j, p) &= A(2, j, p)^2 + B(2, j, p)^2 \\ E_{-1}(1, j, p) &= A_{-1}(1, j, p)^2 + B_{-1}(1, j, p)^2 \\ E_{-1}(2, j, p) &= A_{-1}(2, j, p)^2 + B_{-1}(2, j, p)^2 \\ E'(1, j, p) &= E(1, j, p) - q \cdot E_{-1}(2, j, p) \\ E'(2, j, p) &= E(2, j, p) - q \cdot E(1, j, p) \end{aligned}$$

【0099】

上記〔数式15〕における6つの式のうち、上から5番目（下から2番目）の式は、あるAタイプの音響フレームに着目したときに、その前半の $N/2$ サンプルが重複するBタイプの音響フレームからの残響成分を除去するためのものである。また、上から6番目（下から1番目）の式は、あるBタイプの音響フレームに着目したときに、その前半の $N/2$ サンプルが重複するAタイプの音響フレームからの残響成分を除去するためのものである。 $E'(1, j, p) = 0$ 、 $E'(2, j, p) = 0$ とし、〔数式15〕の上から5番目、6番目の式に従った処理の結果、負の値となった場合には、0に設定する。

【0100】

上記〔数式15〕において、 q は残響成分の大きさを示す係数であるが、この係数 q は1未満の値を持つものであり、実験の結果、 $N = 4096$ のとき $q = 0.06$ 、 $N = 2048$ のとき $q = 0.12$ 、 $N = 1024$ のとき $q = 0.24$ 、 $N = 512$ のとき $q = 0.48$ が最適である。

【0101】

そして、残響成分を除去しない場合の各スペクトル集合の強度値 E_1 、 E_2 、 E_3 、 E_4 、除去した場合の強度値 E_1' 、 E_2' 、 E_3' 、 E_4' を、算出した基本強度値 $E(1, j, p)$ 、 $E(2, j, p)$ 、 $E'(1, j, p)$ 、 $E'(2, j, p)$ を用いて以下の〔数式16〕に従った処理により算出する。

【0102】

〔数式16〕

$$E_1(p) = \sum_{j=m, \dots, m+G-1} E(1, j, p)$$

$$\begin{aligned}
E_2(p) &= \sum_{j=m, \dots, m+G-1} E(2, j, p) \\
E_3(p) &= \sum_{j=m+G, \dots, m+2G-1} E(1, j, p) \\
E_4(p) &= \sum_{j=m+G, \dots, m+2G-1} E(2, j, p) \\
E_1'(p) &= \sum_{j=m, \dots, m+G-1} E'(1, j, p) \\
E_2'(p) &= \sum_{j=m, \dots, m+G-1} E'(2, j, p) \\
E_3'(p) &= \sum_{j=m+G, \dots, m+2G-1} E'(1, j, p) \\
E_4'(p) &= \sum_{j=m+G, \dots, m+2G-1} E'(2, j, p)
\end{aligned}$$

【0103】

この $E_1 \sim E_4$ 、 $E_1' \sim E_4'$ については、過去の対応する成分を用いて更新する処理を行う。具体的には、過去の成分を登録する総和値登録テーブルに記録された $Te(1, p) \sim Te(8, p)$ を用いて以下の〔数式17〕に従った処理により算出する。 10

【0104】

〔数式17〕

$$\begin{aligned}
E_1(p) &= 0.67 E_1(p) + 0.33 Te(1, p) \\
E_2(p) &= 0.67 E_2(p) + 0.33 Te(2, p) \\
E_3(p) &= 0.67 E_3(p) + 0.33 Te(3, p) \\
E_4(p) &= 0.67 E_4(p) + 0.33 Te(4, p) \\
E_1'(p) &= 0.67 E_1'(p) + 0.33 Te(5, p) \\
E_2'(p) &= 0.67 E_2'(p) + 0.33 Te(6, p) \\
E_3'(p) &= 0.67 E_3'(p) + 0.33 Te(7, p) \\
E_4'(p) &= 0.67 E_4'(p) + 0.33 Te(8, p)
\end{aligned}$$

20

【0105】

さらに、 $Te(1, p) \sim Te(8, p)$ は、以下の〔数式18〕に従った処理により更新される。

【0106】

〔数式18〕

$$\begin{aligned}
Te(1, p) &= E_1(p) & Te(2, p) &= E_2(p) \\
Te(3, p) &= E_3(p) & Te(4, p) &= E_4(p) \\
Te(5, p) &= E_1'(p) & Te(6, p) &= E_2'(p) \\
Te(7, p) &= E_3'(p) & Te(8, p) &= E_4'(p)
\end{aligned}$$

30

【0107】

結局〔数式15〕〔数式16〕〔数式17〕〔数式18〕により、過去分を考慮した各スペクトル集合の強度値 E_1 、 E_2 、 E_3 、 E_4 が算出されるとともに、当該各スペクトル集合に対応するタイプの別を問わない直前の音響フレームにおけるスペクトル集合の強度に q を乗じた値を減じることにより強度値 E_1' 、 E_2' 、 E_3' 、 E_4' が算出されることになる。

【0108】

また、符号判定パラメータ算出手段140は、残響成分を除去せずに算出した強度値 E_1 、 E_2 、 E_3 、 E_4 を用いて、以下の〔数式19〕に従った処理を実行し、符号判定パラメータ C を算出する。 40

【0109】

〔数式19〕

1) $E_1(p) > E_2(p)$ かつ $E_4(p) > E_3(p)$ の場合

$B = 0$ と仮判定、

$$C = \{ E_1(p) \cdot E_4(p) - E_2(p) \cdot E_3(p) \} / \{ E_1(p) \cdot E_4(p) + E_2(p) \cdot E_3(p) \}$$

2) $E_2(p) > E_1(p)$ かつ $E_3(p) > E_4(p)$ の場合

$B = 1$ と仮判定、

$$C = \{ E_2(p) \cdot E_3(p) - E_1(p) \cdot E_4(p) \} / \{ E_1(p) \cdot E_4(p) + E_2(p) \cdot E_3(p) \}$$

50

3) $E_1(p) \cdot E_4(p) > E_2(p) \cdot E_3(p)$ の場合

$B = 0$ と仮判定、

$$C = \{ E_1(p) \cdot E_4(p) - E_2(p) \cdot E_3(p) \} / \{ E_1(p) \cdot E_4(p) + E_2(p) \cdot E_3(p) \}$$

4) 上記 1) ~ 3) 以外の場合

$B = 1$ と仮判定、

$$C = \{ E_2(p) \cdot E_3(p) - E_1(p) \cdot E_4(p) \} / \{ E_1(p) \cdot E_4(p) + E_2(p) \cdot E_3(p) \}$$

【0110】

また、符号判定パラメータ算出手段 140 は、残響成分を除去して算出した強度値 E_1' 、 E_2' 、 E_3' 、 E_4' を用いて、以下の〔数式 20〕に従った処理を実行し、補正符号判定パラメータ C' を算出する。

【0111】

〔数式 20〕

1) $E_1'(p) > E_2'(p)$ かつ $E_4'(p) > E_3'(p)$ の場合

$B' = 0$ と仮判定、

$$C' = \{ E_1'(p) \cdot E_4'(p) - E_2'(p) \cdot E_3'(p) \} / \{ E_1'(p) \cdot E_4'(p) + E_2'(p) \cdot E_3'(p) \}$$

2) $E_2'(p) > E_1'(p)$ かつ $E_3'(p) > E_4'(p)$ の場合

$B' = 1$ と仮判定、

$$C' = \{ E_2'(p) \cdot E_3'(p) - E_1'(p) \cdot E_4'(p) \} / \{ E_1'(p) \cdot E_4'(p) + E_2'(p) \cdot E_3'(p) \}$$

3) $E_1'(p) \cdot E_4'(p) > E_2'(p) \cdot E_3'(p)$ の場合

$B' = 0$ と仮判定、

$$C' = \{ E_1'(p) \cdot E_4'(p) - E_2'(p) \cdot E_3'(p) \} / \{ E_1'(p) \cdot E_4'(p) + E_2'(p) \cdot E_3'(p) \}$$

4) 上記 1) ~ 3) 以外の場合

$B' = 1$ と仮判定、

$$C' = \{ E_2'(p) \cdot E_3'(p) - E_1'(p) \cdot E_4'(p) \} / \{ E_1'(p) \cdot E_4'(p) + E_2'(p) \cdot E_3'(p) \}$$

【0112】

符号出力手段 150 は、1つの基準フレームに対応する音響フレーム（基準フレームおよび位相変更フレーム）の中から最適な位相であると判断されるものを判断し、その音響フレームの状態に対応する符号を出力する機能を有している。符号判定パラメータ算出手段 140 と符号出力手段 150 により符号化手段が構成される。付加情報抽出手段 160 は、符号出力手段 150 により出力された 2 値の配列を N_h ビット単位で抽出し、巡回符号表を参照することにより、 N_w ビットの基準符号に変換し、さらに所定の規則により変換して意味のある付加情報として抽出する機能を有している。音響フレーム保持手段 170 は、チャンネルごとの A タイプ、B タイプそれぞれについて連続する 2 個の基準フレーム（チャンネルごとに計 4 個の基準フレーム）を保持可能なバッファメモリである。変換テーブル作成手段 180 は、図 2 に示した変換テーブル作成手段 80 と同様、 N_w ビットが取り得る 2 の N_w 乗個の全ての基準符号に対して、互いにハミング距離が少なくとも 2 以上となる N_h ($> N_w$) ビットのハミング符号を割り当てることにより、 N_w ビットの基準符号と N_h ビットのハミング符号が対応付けられた符号変換テーブルを作成する機能を有している。

【0113】

図 13 に示した各構成手段は、現実には情報処理機能を有する小型のコンピュータおよびその周辺機器等のハードウェアに専用のプログラムを搭載することにより実現される。特に、本発明の目的をより簡易に達成するためには、プログラムを実行するコンピュータとしての機能を備えた携帯電話機等の携帯型端末装置をハードウェアとして用いることが

10

20

30

40

50

望ましい。

【 0 1 1 4 】

(3 . 2 . 抽出装置の処理動作)

次に、図 1 3 に示した音響信号からの情報の抽出装置の処理動作について、図 1 2 のフローチャートに従って説明する。抽出装置を起動すると、まず、本装置では、位相判定テーブル $S(p)$ 、位相決定ログ、位相確定フラグ、総和値累積テーブル $Te(1, p) \sim Te(8, p)$ 、ビットカウンタを初期化する ($S200$)。位相判定テーブル $S(p)$ は、位相を判定するためのテーブルであり、 p は $0 \sim 5$ の整数値をとる。初期値は $S(p) = 0$ に設定されている。位相決定ログは、1つの基準フレームと5つの位相変更フレームの組ごとに、決定された位相すなわち位相番号 p を記録していくものであり、初期状態では 0 が設定されている。位相確定フラグは、位相が確定しているかどうかを示すフラグであり、初期状態では Off に設定されている。総和値累積テーブル $Te(1, p) \sim Te(8, p)$ は、過去 (バイト周期 $\times N_h$ フレーム) 分の総和値 $E1 \sim E4$ 、残響補正後の総和値 $E1' \sim E4'$ を6通りの位相別に保存したものであり、初期状態では $Te(1, p) \sim Te(8, p) = 0$ に設定されている。ビットカウンタについては初期値として 0 を設定する。

【 0 1 1 5 】

続いて、変換テーブル作成手段 180 が、巡回符号表を作成する ($S210$)。巡回符号表とは、図 7 に示した符号変換テーブルにおいて、変換元の 7 ビット基準符号に対応する変換後の 12 ビットハフマン符号のビット列を 1 ビットずつ移動させた巡回ビットパターンを記録したものである。変換テーブル作成手段 180 は、まず、変換テーブル作成手段 80 と同様、図 6 のフローチャートに従った処理により図 7 に示したような符号変換テーブルを作成した後、巡回符号表を作成する。具体的には、変換テーブル作成手段 180 は、図 7 に示したような符号変換テーブルにおいて、7 ビット基準符号に対応付けられた 12 ビットハミング符号のビット列を 1 ビットずつ移動した巡回ビットパターンを作成し、7 ビット基準符号に対応付けて記録する。例えば、図 7 の 7 ビット基準符号 “0” には、12 ビットハミング符号 “0000, 0000, 0001” が正規のビットパターンとして対応付けられているので、変換テーブル作成手段 180 は、1 ビット移動した 12 ビット符号 “0000, 0000, 0010” (10 進表記 “2”)、2 ビット移動した 12 ビット符号 “0000, 0000, 0100” (10 進表記 “4”)、...、11 ビット移動した 12 ビット符号 “1000, 0000, 0000” (10 進表記 “2048”) までの 11 個の 12 ビット符号を巡回ビットパターンとして作成する。このようにして、作成された巡回符号表の一例を図 1 5 に示す。

【 0 1 1 6 】

図 1 5 の巡回符号表においては、いずれも 10 進表記で示してある。例えば、1 行目の 7 ビット基準符号 “0” に対しては、正規の 12 ビットハミング符号 “1” と 11 個の巡回ビットパターン “2” “4” “8” “16” “32” “64” “128” “256” “512” “1024” “2048” が対応付けて記録される。

【 0 1 1 7 】

このように、初期値が設定され、巡回符号表が作成された状態で、携帯端末に対して、抽出装置としての起動の指示を行う。これは、例えば、抽出装置を携帯電話機等の携帯端末で実現している場合は、所定のボタンを操作することにより実行できる。抽出装置は、指示が入力されると、音響信号入力手段 100 が、流れている音楽を録音し、デジタル化してデジタル音響信号として入力する。具体的には、無指向性マイクロフォン (または指向性マイクロフォンの一方のチャンネル) から入力される音声を、A/D 変換器によりデジタル化する処理を行うことになる。

【 0 1 1 8 】

続いて、基準フレーム獲得手段 110 が、音響信号入力手段 100 から入力された音響信号から、所定数のサンプルで構成される音響フレームを基準フレームとして抽出する ($S201$)。具体的には、A タイプ、B タイプについての基準フレームを抽出して音響フ

10

20

30

40

50

レーム保持手段 170 に読み込むことになる。基準フレーム獲得手段 110 が基準フレームとして読み込む 1 音響フレームのサンプル数は、図 2 に示した音響フレーム読込手段 10 で設定されたものと同じにする必要がある。したがって、本実施形態の場合、基準フレーム獲得手段 110 は、A タイプ、B タイプについてそれぞれ 4096 サンプルずつ、順次基準フレームとして読み込んでいくことになる。音響フレーム保持手段 170 には、各チャンネルについて、A タイプ、B タイプ 2 個ずつの基準フレーム、すなわち 2.5 N サンプルが格納可能となっており、新しい基準フレームが読み込まれると、古い基準フレームを破棄するようになっている。したがって、音響フレーム保持手段 170 には、常に基準フレーム 4 個分（連続する 10240 サンプル）が格納されていることになる。

【0119】

10

抽出装置で処理する音響フレームは、先頭から途切れることなく隣接して設定される基準フレームと、この基準フレームと位相を変更した位相変更フレームとに分けることができる。基準フレームについては、A タイプの音響フレームと B タイプの音響フレームが、2048 サンプルずつ重複している場合、最初の基準フレームとしてサンプル番号 1 からサンプル番号 4096 までを設定したら、次の基準フレームは、サンプル番号 2049 からサンプル番号 6144、さらに次の基準フレームは、サンプル番号 4097 からサンプル番号 8192、さらに次の基準フレームは、サンプル番号 6145 からサンプル番号 10240、というように途切れることなく設定される。そして、各基準フレームについて、1/6 フレーム（約 683 サンプル）ずつ移動した 5 個の位相変更フレームを設定する。例えば、最初の基準フレームについては、サンプル番号 683、1366、2049、2732、3413 から始まる 4096 のサンプルで構成される 5 個の位相変更フレームが設定されることになる。続いて、周波数変換手段 130、符号判定パラメータ算出手段 140 が、読み込んだ各音響フレームについて、位相を特定した後、埋め込まれている情報を判定し、対応する符号を出力する（S202）。出力される情報の形式は、付加情報が埋め込まれる場合に対応する値 1、値 2 の 2 通りの形式となる。

20

【0120】

ここで、ステップ S202 の位相確定および符号判定の詳細を図 16 のフローチャートに従って説明する。まず、位相確定フラグが On であるか Off であるかの確認を行う（S301）。位相確定フラグが On である場合は、位相確定処理（S303～S309）を行わず、符号判定処理のみを行う（S302）。ただし、初期状態では位相は確定しておらず、位相確定フラグが Off となっているので、候補符号テーブル B(p) の初期化を行う（S303）。候補符号テーブル B(p) は、1 つの基準フレームおよび 5 個の位相変更フレームを特定する $p = 0 \sim 5$ の位相番号および、この 6 個の音響フレームの状態から得られる 2 値の符号を記録するものである。

30

【0121】

続いて、符号判定パラメータ算出手段 140 は、符号判定処理を行う（S302）。ここで、符号判定処理の詳細を図 17 に示す。まず、周波数変換手段 130 が、読み込んだ各音響フレームに対して、周波数変換を行って各窓スペクトルを得る（S401）。具体的には、上記〔数式 12〕〔数式 13〕に従った処理を実行し、変換データの実部 $A(1, j, p)$ 、虚部 $B(1, j, p)$ 、実部 $A(2, j, p)$ 、虚部 $B(2, j, p)$ を得る。

40

【0122】

S401 の処理については、実際には、A タイプ音響フレームについて周波数変換を行った後、 $N/2$ サンプルだけシフトして B タイプ音響フレームについて周波数変換を行う。これらの変換データ $A(2, j, p)$ 、 $B(2, j, p)$ 、 $A(2, j, p)$ 、 $B(2, j, p)$ に対して、上記〔数式 14〕に従った処理を実行し、振幅変換を行う（S402）。

【0123】

上記周波数変換手段 130 における処理により、周波数に対応した成分であるスペクトルで表現されたフレームスペクトルが得られる。続いて、符号判定パラメータ算出手段 1

50

40は、符号判定パラメータ C 、 C' を上述のようにして算出した後、符号判定パラメータ C 、 C' を用いて、変更対象周波数帯の成分の状態がどのような状態であるか、すなわち、1ビットの値としてどのような値が埋め込まれていたかを判断する処理を行う(S403)。具体的には、上記〔数式15〕～〔数式19〕に従った処理を実行して、符号判定パラメータ C 、 C' を算出する。そして、この両者を比較して、 $C \geq C'$ であれば候補符号 B を候補符号テーブル $B(p)$ に設定し、 $C < C'$ であれば候補符号 B' を候補符号テーブル $B(p)$ に出力する。

【0124】

S403の具体的な処理手順としては、まず、上記〔数式15〕の1番目の式と2番目の式を利用して基本強度値 $E(1, j, p)$ 、 $E(2, j, p)$ を算出した後、上記〔数式16〕の1番目から4番目までの式を利用して $E1(p)$ 、 $E2(p)$ 、 $E3(p)$ 、 $E4(p)$ を算出し、上記〔数式19〕に従って候補符号 B および符号判定パラメータ C を算出する処理を行う。続いて、上記〔数式15〕の3番目から6番目までの式を利用して、残響補正処理を行って基本強度値 $E'(1, j, p)$ 、 $E'(2, j, p)$ を算出する。そして、上記〔数式16〕の5番目から8番目までの式を利用して $E1'(p)$ 、 $E2'(p)$ 、 $E3'(p)$ 、 $E4'(p)$ を算出し、上記〔数式20〕に従って候補符号 B' および補正符号判定パラメータ C' を算出する処理を行う。

【0125】

また、上記判定の結果、位相 p における候補符号テーブル $B(p)$ に値1、値2のいずれかを出力した場合には、さらに、以下の〔数式21〕に従って位相判定テーブル $S(p)$ の更新を行う(S404)。

【0126】

〔数式21〕

$C \geq C'$ である場合、 $S(p) \leftarrow S(p) + C$

$C < C'$ である場合、 $S(p) \leftarrow S(p) + C'$

【0127】

ここで、図16のフローチャートに戻って、符号判定パラメータ算出手段140は、候補符号テーブル $B(p)$ に、位相 p において符号判定処理(S302)で仮決定された符号を保存する(S304)。

【0128】

続いて、全ての位相番号 p に対応する処理を終えたかどうかを判定する(S305)。これは、ある基準フレームに対して全ての位相変更フレームの処理を行ったかどうかを判定している。本実施形態では、 p が0～5までの値をとるので、6回分処理していない場合は、処理していた音響フレームから所定サンプル数ずらして、位相の異なる音響フレームを設定し、S302に戻って処理を繰り返す。なお、 $p = 0$ の場合が基準フレームであり、 $p = 1 \sim 5$ の場合が位相変更フレームである。全ての位相番号 p に対応する処理を終えた場合は、位相判定テーブル $S(p)$ の値が最大となる位相番号 p_{max} に対応する位相が最適位相であると判定し、候補符号テーブル $B(p)$ に記録されている符号 $B(p_{max})$ を出力する(S306)。

【0129】

続いて、位相決定ログの更新を行う(S307)。位相決定ログとは、1つの基準フレームと5つの位相変更フレームの組ごとに、決定された位相すなわち位相番号 p を記録するものである。そして、位相決定ログを参照して、位相が過去所定回数同一であったかどうかを判断する(S308)。本実施形態では、この回数を10回としている。位相が過去所定回数同一であった場合には、位相確定フラグをOnに設定する(S309)。これにより、同一位相が所定回数続いた場合には、最適位相は p_{max} となる可能性が高いため、位相確定処理(S303～S309)を行わず、位相番号 $p = p_{max}$ に対してのみ符号判定処理(S302)を行うことになる。

【0130】

再び図14のフローチャートに戻って説明する。位相確定および符号出力が行なわれた

ら、出力された符号の値に対応する 1 ビットをバッファに保存する (S 2 0 3)。次に、ビットカウンタを “ 1 ” だけカウントアップする (S 2 0 4)。

【 0 1 3 1 】

次に、ビットカウンタが 1 1 以下であるか 1 2 以上であるかを判断する (S 2 0 5)。ビットカウンタが 1 1 以下の場合は、 S 2 0 1 に戻って、次の A タイプ、 B タイプの基準フレームを抽出する処理を行う。

【 0 1 3 2 】

ビットカウンタが 1 2 以上である場合は、バッファに保存された 1 2 ビットのビット配列と巡回符号表内のビットパターンとの照合を行う (S 2 0 6)。照合の結果、一致するビットパターンが 1 2 ビットハミング符号または巡回ビットパターンとして存在する場合には、適合パターン有りと判断し、対応する基準符号 7 ビットを抽出する (S 2 1 1)。音響フレームから抽出された 1 2 ビットのビット列が巡回ビットパターンのいずれかであった場合、これらのビットパターンは埋め込み時に埋め込んでいるはずであるので、対応する 1 2 ビットハミング符号のビットパターンをずれた位置で読み込んでいると判断し、7 ビット基準符号を抽出するのである。 S 2 0 6 の照合の結果、一致するビットパターンが 1 2 ビットハミング符号、巡回ビットパターンのいずれにも存在しない場合には、適合パターン無しと判断し、 S 2 0 1 に戻って、次の基準フレームを抽出する処理を行う。

【 0 1 3 3 】

S 2 1 1 において対応する基準符号 7 ビットが抽出された場合は、付加情報抽出手段 1 6 0 が、抽出された 7 ビット基準符号に 1 ビット付加して出力する (S 2 0 8)。具体的には、ビット “ 0 ” を付加することにより A S C I I コードにおける 1 ワードとして出力する。

【 0 1 3 4 】

S 2 0 8 において、7 ビット基準符号に 1 ビット付加して出力した場合には、ビットカウンタを 0 に初期化する (S 2 0 9)。そして、 S 2 0 1 に戻って、次の基準フレームから抽出する処理を行う。本発明では、埋め込み時に N a 回重複して同一ワードを埋め込んでいるが、埋め込み時の先頭ビットから 1 2 ビットが偶然にも抽出されない限り、通常 (N a - 1) 回の同一ワードが抽出されることになる (1 ワード分は前後の異なるワード間をまたがって抽出されるため、 S 2 0 6 の照合の結果、通常エラーになる)。この場合、実際に情報として必要なのは 1 ワード分のみであるため、後続する (N a - 2) ワード分の 7 ビット基準符号を読み飛ばすのである。もちろん、状況によっては、 (N a - 2) ワード以下の 7 ビット基準符号しか読み込まれない場合もあれば、全く読み込まれない場合もある。

【 0 1 3 5 】

図 1 4 に示す処理を各基準フレームに対して実行することにより、付加情報が抽出されることになる。 S 2 0 1 において全ての基準フレームが抽出されたと判断された場合には、処理を終了する。

【 0 1 3 6 】

上記 S 2 0 8 の処理において、付加情報抽出手段 1 6 0 は、まず、符号出力手段 1 5 0 により出力された値を、所定の規則により変換して意味のある付加情報として抽出する。所定の規則としては、情報を埋め込む者が意図した情報が受け取った者に認識可能な状態にすることができるものであれば、さまざまな規則が採用できるが、本実施形態では、 A S C I I コードを採用している。すなわち、付加情報抽出手段 1 6 0 は、符号判定パラメータ算出手段 1 4 0 が判定し、符号出力手段 1 5 0 から出力される符号から得られるビット値の配列を 1 バイト (8 ビット) 単位で認識し、これを A S C I I コードに従って文字情報を認識する。このようにして得られた文字情報は、表示装置 (図示省略) の画面に表示出力される。

【 0 1 3 7 】

従って、埋め込み装置により音響信号に、その楽曲の曲名やアーティスト等の属性情報を文字情報として埋め込んでおけば、利用者は、その音楽が流れているのを聞いて、その曲

10

20

30

40

50

名やアーティストを知りたいと思ったときに、抽出装置として機能する自身の携帯端末に所定の操作を行えば、自身の携帯端末の画面に曲名やアーティスト等の属性情報が文字情報として表示されることになる。

【0138】

(3.3.位相補正処理について)

上記のように、抽出時には、埋め込み時に埋め込んだ音響フレームに対応して、音響信号を読み込むことができるとは限らない。そこで、音響フレームの位相をずらして複数通り(本実施形態では6通り)で読み込み、その中で最適な位相を決定し、その位相で特定される音響フレームに対応する符号を出力することになっている。例えば、6通りで読み込む場合、先頭の音響フレームは、本来サンプル番号1~4096のサンプルであるが、サンプル番号1、683、1366、2049、2732、3413から始まる4096のサンプルで構成される6個の各音響フレームに対して処理を行い、最適な音響フレームに対応する符号を出力することになる。なお、図16のフローチャートを用いて説明したように、本実施形態では、同一位相が所定回数連続した場合には、それ以降はその位相を確定したものとして処理を行う。

10

【0139】

(4.ビット配列のビット数 N_h 、ハミング距離を変更した場合)

上記実施形態では、ビット配列作成の際、作成されるビット配列は、12ビット(N_h ビット)でハミング距離が4となるようにしたが、ビット配列のビット数 N_h 、およびハミング距離は、適宜変更することが可能である。一般に、ビット配列のビット数 N_h を大きくすれば、符号化効率は悪くなるが、訂正可能な誤りビット数は増えることになる。したがって、マイクロフォンの感度や演算処理能力が比較的高い抽出装置を用いる場合には、ビット数、ハミング距離が小さいビット配列を作成するようにし、マイクロフォンの感度や演算処理能力が比較的低い抽出装置を用いる場合には、ビット数、ハミング距離が大きいビット配列を作成するようにする。

20

【0140】

ここでは、ビット配列のビット数 N_h を16、ハミング距離を6とした場合の例について説明する。ビット配列のビット数を16、ハミング距離を6とした場合も、変換テーブル作成手段80による符号変換テーブル作成は、基本的には、図6に示したフローチャートに従って行われる。ただし、HFが16ビット符号である点、S603において16ビット符号HFと符号変換テーブルに既に登録されている他の16ビット符号とのハミング距離を算出した後、S604、S605のいずれに進むかの閾値が、ハミング距離が6未満か、6以上かで決まる点が、ビット数 $N_h=12$ 、ハミング距離4の場合と異なっている。

30

【0141】

このようにして作成された符号変換テーブルを図18に示す。図18に示すように、符号変換テーブルには、7ビット符号が取り得る各値0~127について、16ビット符号が対応付けて登録されている。16ビット符号の2進表記を見るとわかるように、全ての16ビット符号は互いにハミング距離が6以上となっている。なお、図18の例では、説明の便宜上、7ビット符号の10進表記と、16ビット符号の10進表記及び2進表記を示しているが、現実には、7ビット符号のビット列と16ビット符号のビット列が対応付けて登録されている。

40

【0142】

符号変換テーブル作成処理後は、図8のフローチャートに従って埋め込み処理が行われる。この際、1ワードとして、16ビット分の処理が行われることになる。

【0143】

抽出時には、変換テーブル作成手段180が、最初に符号変換テーブルを作成した後、巡回符号表を作成する。巡回符号表が作成された後は、図14、図16、図17に従った処理により、付加情報が抽出されることになる。16ビット符号を扱う場合、図14におけるS205においては、ビットカウンタが15以下か16以上かで判断を行うことにな

50

る。

【0144】

16ビット符号についての巡回符号表を作成する場合、変換テーブル作成手段180は、図18に示したような符号変換テーブルにおいて、7ビット基準符号に対応付けられた16ビットハミング符号のビット列を1ビットずつ移動した巡回ビットパターンを作成し、7ビット基準符号に対応付けて記録する。例えば、図18の7ビット基準符号“0”には、16ビットハミング符号“0000, 0000, 0000, 0001”が正規のビットパターンとして対応付けられているので、変換テーブル作成手段180は、1ビット移動した16ビット符号“0000, 0000, 0000, 0010”(10進表記“2”)、2ビット移動した12ビット符号“0000, 0000, 0000, 0100”(10進表記“4”)、...、15ビット移動した16ビット符号“1000, 0000, 0000, 0000”(10進表記“32768”)までの15個のビットパターンを巡回ビットパターンとして作成する。このようにして、作成された巡回符号表の一例を図19に示す。ビット配列のビット数Nhを16、ハミング距離を6とした場合図19に示した巡回符号表を用いてビットパターンの照合を行うことになる。

10

【0145】

ところが、図18に示した巡回符号表には図示されていない基準符号の117、125、126(ハミング符号では62713、65132、65428)に問題があり、これら3種の符号は他のハミング距離とは6ビット離れているが、巡回符号(巡回ビットパターン)を含めるとハミング距離が0、即ち重複してしまう場合が生じる。即ち、これら3種の符号を含めて埋め込まれると、巡回ビットパターンから基準符号を特定できず、文字化けを起こす可能性がある。このような問題を回避するため、以下、巡回符号に基づいた改良型16ビットハミング符号の作成方法について、図20のフローチャートに基づいて述べる。

20

【0146】

(5.巡回符号を考慮して、同ビット数Nhのハミング符号を再作成した場合)

最初に初期化処理を行う(S611)。具体的には、基準符号“0”に16ビット符号“1”を対応付けて符号変換テーブルのi(=0)番目に登録するとともに、基準符号KFの初期値を1、16ビット符号HFの初期値を2とする。基準符号のビット数は図6の場合と異なり7ビットの範囲を若干超えるため、8ビットで設定する。続いて、i=0に初期設定する(S612)。次に、16ビット符号HF及びその15通りの巡回符号と、符号変換テーブルのi番目に既に登録されている他の16ビット符号及びその15通り巡回符号とのハミング距離を算出する(S613)。この段階で16通り×16通りで256通りのハミング距離が算出される。

30

【0147】

算出したハミング距離の少なくとも1つが3未満の場合、HFの値を1増加して更新した後(S614)、S612に戻って、次の16ビット符号HFについての処理を行う。一方、算出したハミング距離が全て3以上の場合、iの値を1増加して更新する(S615)。そして、iがKF-1未満である場合には、S613に戻って、符号変換テーブル内のi番目に登録されている16ビット符号及びその15通り巡回符号とのハミング距離を算出する。S615の後、iがKF-1以上となった場合には、16ビット符号HFを符号変換テーブルの7ビット符号KFの位置に登録するとともに、KF、HFの値をそれぞれ1ずつ増加して更新する(S616)。そして、HFが65536未満である場合は、S612に戻って、次の8ビット符号KFについての処理を行う。HFが65536以上である場合は、16ビットの範囲で表現可能な全ての16ビット符号HFが登録されたことになるので、符号変換テーブル作成処理を終了する。

40

【0148】

このようにして作成された符号変換テーブルを図21に示す。図21は図7や図18に対応させるために、巡回符号を考慮しないハミング符号との対応部分だけを抜き出したものである。この符号変換テーブルには、16ビットのハミング符号が定義可能な0~65

50

5 3 5 の範囲内で、各々 1 6 通りの巡回符号が互いにハミング距離 3 以上になるように、8 ビットの基準符号を定義したもので、基準符号としては 7 ビットより若干多い 0 ~ 1 3 2 の値が表現可能となっている。

【0 1 4 9】

変換テーブル作成手段 1 8 0 は、図 2 1 に示したような符号変換テーブルにおいて、8 ビット基準符号に対応付けられた 1 6 ビットハミング符号のビット列を 1 ビットずつ移動した巡回ビットパターンを作成し、8 ビット基準符号に対応付けて記録する。このようにして、図 2 2 に示すような巡回符号表を作成する。

【0 1 5 0】

(6 . 信号成分が小さい場合にさらに確実に情報の埋め込みを行う手法)

10

ここまでの説明のように、本発明では、元の信号成分に無音に近い部分があっても、情報の埋め込みを可能とするものである。このままでも、もちろん十分であるが、本発明においては、より確実に情報を埋め込むような処理を付加することも可能である。具体的には、周波数変換を行う前に、音響信号に聴取困難な微弱な白色ノイズを重畳させる。

【0 1 5 1】

この場合、図 2 に示した音響フレーム読込手段 1 0 は、所定数のサンプルを 1 音響フレームとして読み込んだ後、所定の振幅範囲で所定サンプル数単位で同一値を加算または減算する機能を有している。本実施形態では、左チャンネル信号 $X_l(i)$ 、右チャンネル信号 $X_r(i)$ ($i = 0, \dots, N - 1$) に対して、 $-32 \sim +32$ の振幅範囲 ($X_l(i)$ 、 $X_r(i)$ のとり得る値が $-32768 \sim +32767$ の場合) で一様乱数 H を発生させ、 K (本実施形態では $K = 5$) 個のサンプル間は同一値 H を用いて以下の〔数式 2 2〕に示すような処理を実行し、 $X_l(i)$ 、 $X_r(i)$ の値を更新する。

20

【0 1 5 2】

〔数式 2 2〕

$$X_l(i) \leftarrow X_l(i) + H$$

$$X_r(i) \leftarrow X_r(i) - H$$

【0 1 5 3】

すなわち、上記〔数式 2 2〕に従った処理を N サンプルに渡って実行することにより、1 つの音響フレームに白色ノイズが発生されることになる。この白色ノイズの発生処理は、図 8 のフローチャートにおいては、S 1 0 3、S 1 0 4 の音響フレーム抽出処理の直後に行われることになる。

30

【0 1 5 4】

上記のように、白色ノイズが発生させた場合であっても、抽出側の、音響信号からの情報の抽出装置の構成は図 1 3 と同一であり、処理動作は図 1 4、図 1 6、図 1 7 のフローチャートに従ったものと同一である。

【0 1 5 5】

(7 . ネットワーク接続への応用)

本発明は、特に、抽出した情報を利用して、コンピュータネットワークに接続する場合に応用することができる。例えば、ネットワーク接続装置内に本発明の抽出装置を組み込んでおき、抽出した情報を URL 等のコンピュータにアクセスするための情報に変換して利用する。このような用途に用いる場合は、URL 等に変換するテーブルさえ用意しておけば、埋め込んでおく情報は、非常に短いもので済む。したがって、利用者が録音するのに面倒とならない極めて短時間であっても、その間に、繰り返し埋め込んでおくことができる。

40

【0 1 5 6】

(8 . その他)

以上、本発明の好適な実施形態について説明したが、本発明は上記実施形態に限定されず、種々の変形が可能である。例えば、上記実施形態では、1 音響フレームのサンプル数 $N = 4096$ としたが、 $N = 2048$ 、 1024 、 512 等を設定するようにしても良い。これにより、同一時間あたりの音響フレーム数が、2 倍、4 倍、8 倍となり、全体とし

50

て 2 ～ 8 倍の情報の埋め込みが可能となる。

【 0 1 5 7 】

また、上記実施形態では、音響信号からの情報の抽出装置を、携帯電話機等の携帯型端末装置単体で実現した場合を例にとって説明したが、他のコンピュータと連携して実現するようにしても良い。具体的には、携帯型端末装置と専用コンピュータを無線通信可能に接続し、音響信号入力手段 1 0 0 ～ 音響フレーム保持手段 1 7 0 の構成要素のうち、演算負荷の大きいものは、専用コンピュータで処理する。例えば、音響信号入力手段 1 0 0、基準フレーム獲得手段 1 1 0、位相変更フレーム設定手段 1 2 0、付加情報抽出手段 1 6 0、変換テーブル作成手段 1 8 0 を携帯型端末装置に備え、周波数変換手段 1 3 0、符号判定パラメータ算出手段 1 4 0、符号出力手段 1 5 0、音響フレーム保持手段 1 7 0 を専用コンピュータに備えるようにして、必要な情報を両者間で通信するようにする。これにより、携帯型端末装置の処理性能が低い場合であっても高速な処理を行うことが可能となる。

10

【 0 1 5 8 】

また、上記実施形態では、2チャンネルのステレオ音響信号を利用した場合を例にとって説明したが、1チャンネルのモノラル音響信号を利用しても良い。この場合は、上記 L チャンネルまたは R チャンネルのいずれか一方に対して行った処理を実行すれば良い。

【 0 1 5 9 】

また、周波数成分変更手段 3 0 による所定周波数成分の変更を、図 5 および〔数式 8〕〔数式 9〕に従って実行するようにしたが、埋め込むべきビット値に応じた周波数成分の変更は、様々なものを用いることができ、例えば特許文献 1 に示したような手法を用いることも可能である。この場合、抽出側においても、当然その埋め込み手法に応じた手法で抽出が行われる。

20

【図面の簡単な説明】

【 0 1 6 0 】

【図 1】付加情報を構成する 1 ワード (1 2 ビット) を 2 回繰り返して埋め込んだ状態を示す図である。

【図 2】本発明に係る音響信号に対する情報の埋め込み装置の機能ブロック図である。

【図 3】本発明で用いる時間方向窓関数を示す図である。

【図 4】人間の聴覚心理特性である音脈分凝の原理の説明図である。

30

【図 5】本発明の一実施形態における変更対象周波数帯の成分の変化の状態を示す図である。

【図 6】符号変換テーブル作成の処理概要を示すフローチャートである。

【図 7】7 ビット符号と 1 2 ビット符号の符号変換テーブルの一例を示す図である。

【図 8】図 2 に示した装置の処理概要を示すフローチャートである。

【図 9】振幅変換を用いない埋め込み処理を説明するための図である。

【図 1 0】振幅変換を用いない抽出処理を説明するための図である。

【図 1 1】振幅変換を用いた埋め込み処理を説明するための図である。

【図 1 2】振幅変換を用いた抽出処理を説明するための図である。

【図 1 3】本発明に係る音響信号からの情報の抽出装置の機能ブロック図である。

40

【図 1 4】図 1 3 に示した装置の処理概要を示すフローチャートである。

【図 1 5】本発明の音響信号からの情報の抽出装置が用いる巡回符号表の一例を示す図である。

【図 1 6】図 1 4 の S 2 0 2 の位相確定および符号出力の詳細を示すフローチャートである。

【図 1 7】図 1 6 の S 3 0 2 の符号判定処理の詳細を示すフローチャートである。

【図 1 8】7 ビット符号と 1 6 ビット符号の符号変換テーブルの一例を示す図である。

【図 1 9】7 ビット符号と 1 6 ビット符号の変換の場合の巡回符号表の一例を示す図である。

【図 2 0】巡回符号に基づいた改良型 1 6 ビットハミング符号に対応した符号変換テーブ

50

ル作成の処理概要を示すフローチャートである。

【図 2 1】8 ビット符号と改良型 1 6 ビット符号の符号変換テーブルの一例を示す図である。

【図 2 2】8 ビット符号と改良型 1 6 ビット符号の変換の場合の巡回符号表の一例を示す図である。

【符号の説明】

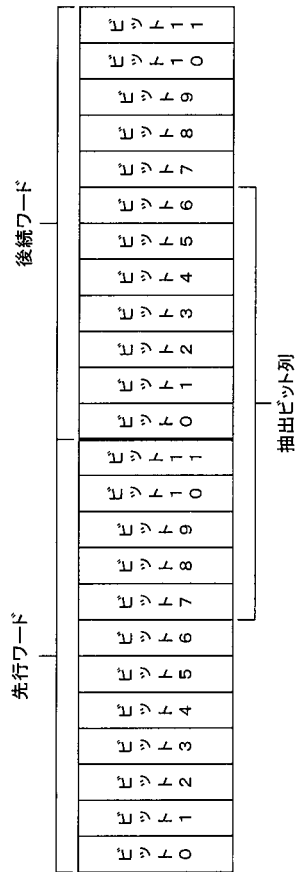
【 0 1 6 1 】

- 1 0 . . . 音響フレーム読込手段
- 2 0 . . . 周波数変換手段
- 3 0 . . . 周波数成分変更手段
- 4 0 . . . 周波数逆変換手段
- 5 0 . . . 改変音響フレーム出力手段
- 6 0 . . . 記憶手段
- 6 1 . . . 音響信号記憶部
- 6 2 . . . 付加情報記憶部
- 6 3 . . . 改変音響信号記憶部
- 7 0 . . . ビット配列作成手段
- 8 0 . . . 変換テーブル作成手段
- 1 0 0 . . . 音響信号入力手段
- 1 1 0 . . . 基準フレーム獲得手段
- 1 2 0 . . . 位相変更フレーム設定手段
- 1 3 0 . . . 周波数変換手段
- 1 4 0 . . . 符号判定パラメータ算出手段
- 1 5 0 . . . 符号出力手段
- 1 6 0 . . . 付加情報抽出手段
- 1 7 0 . . . 音響フレーム保持手段
- 1 8 0 . . . 変換テーブル作成手段

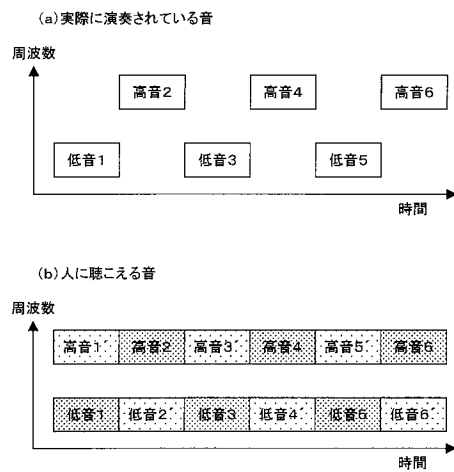
10

20

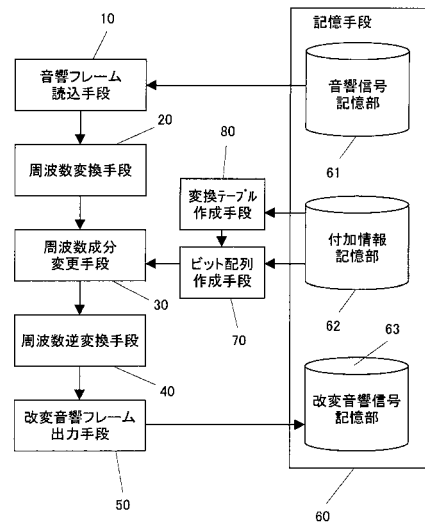
【図 1】



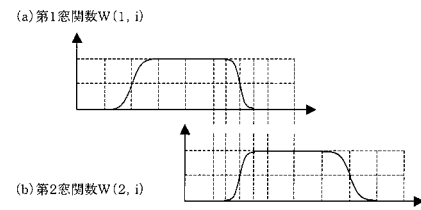
【図 4】



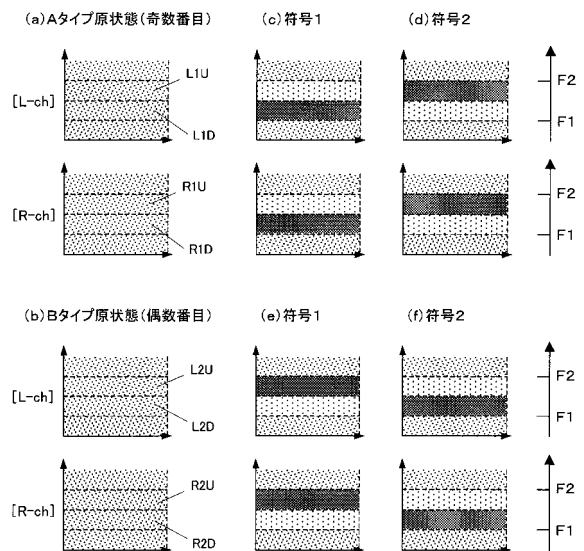
【図 2】



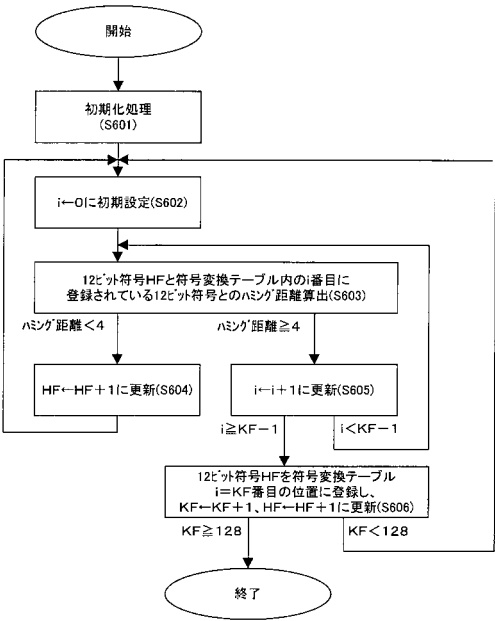
【図 3】



【図 5】



【 図 6 】

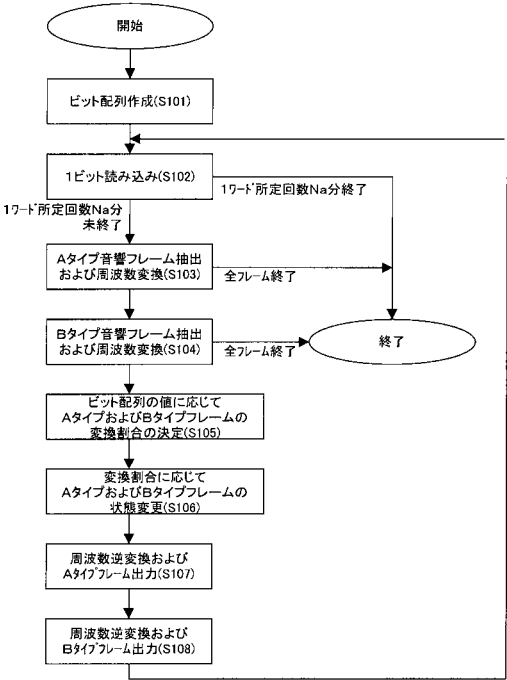


【 図 7 】

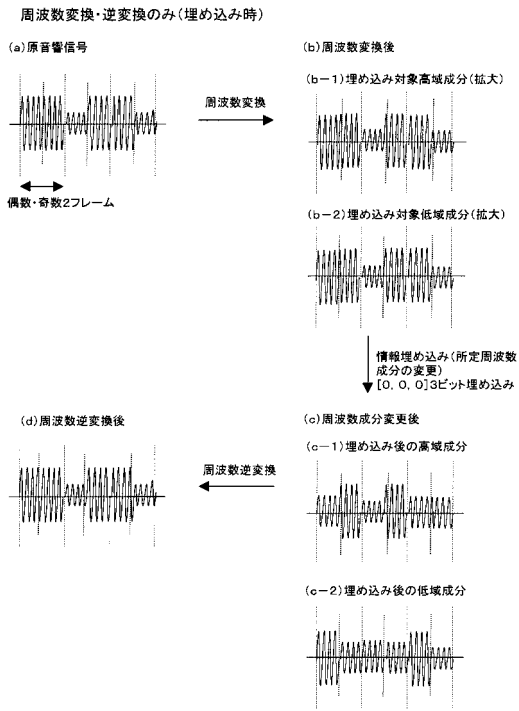
符号変換テーブル

7ビット符号	12ビット符号	
10進表記	10進表記	2進表記
0	1	0000 0000 0001
1	14	0000 0000 1110
2	50	0000 0011 0010
3	61	0000 0011 1101
4	84	0000 0101 0100
5	91	0000 0101 1011
⋮	⋮	⋮
124	4034	1111 1100 0010
125	4045	1111 1100 1101
126	4081	1111 1111 0001
127	4094	1111 1111 1110

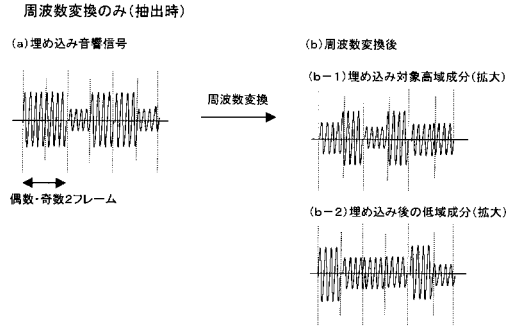
【 図 8 】



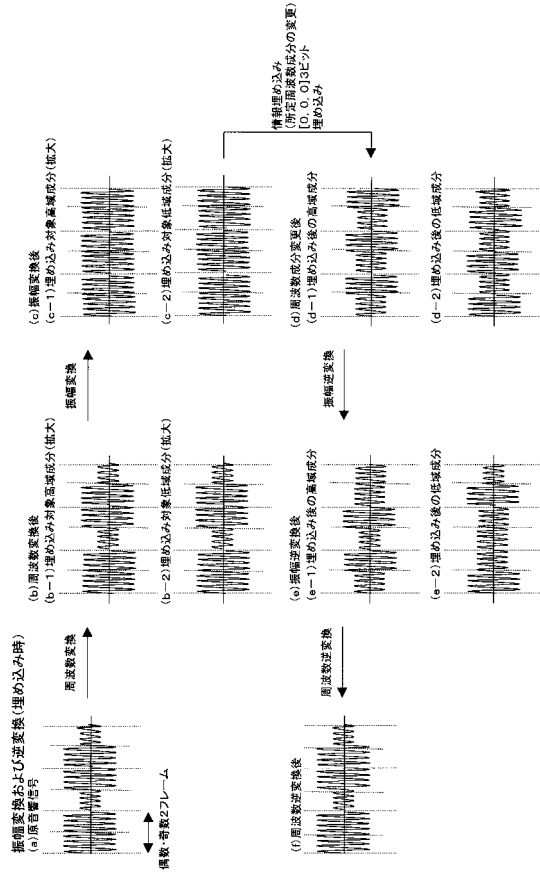
【 図 9 】



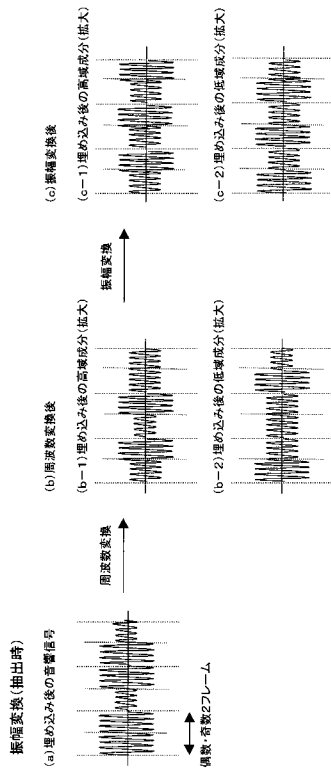
【図 10】



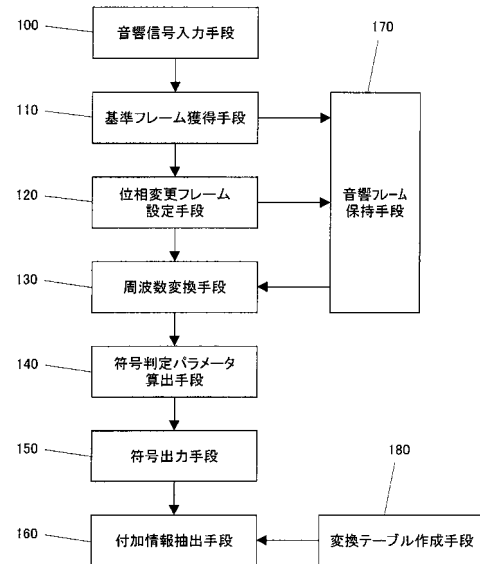
【図 11】



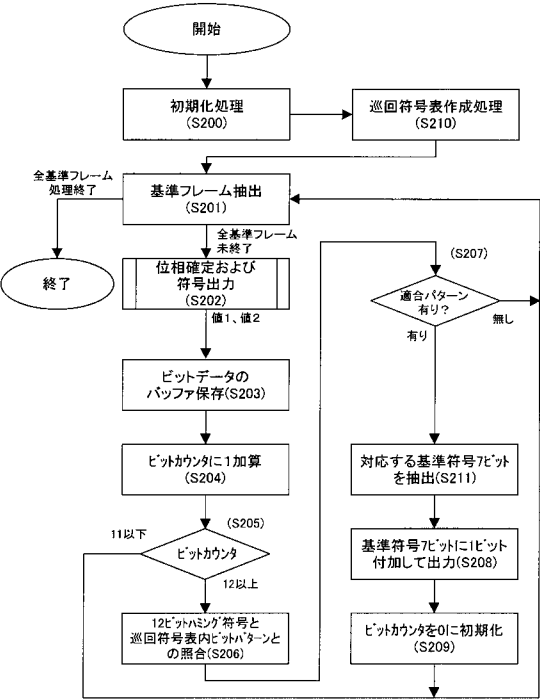
【図 12】



【図 13】



【 図 1 4 】

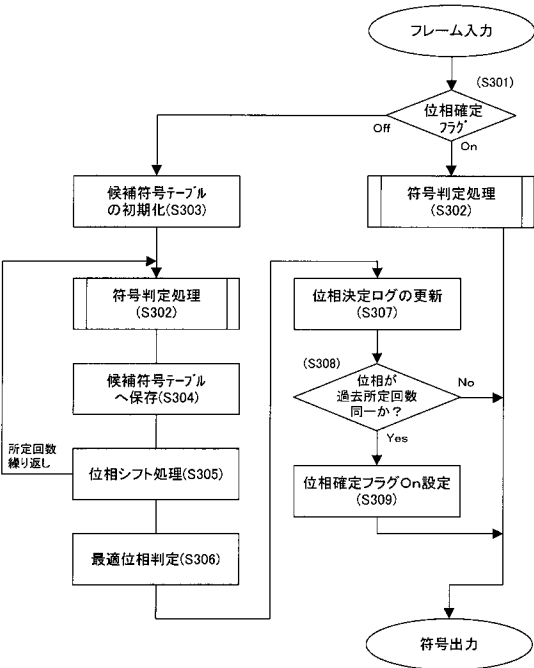


【 図 1 5 】

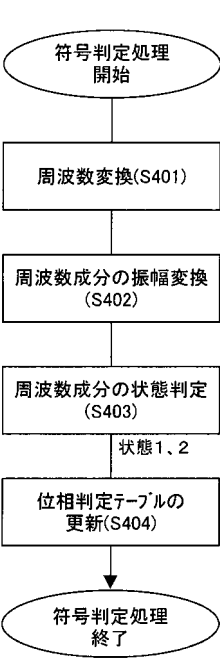
巡回符号表

7ビット符号 (基準符号)	12ビット符号 (ハミング符号)	巡回ビットパターン(11種類)															
10進表記	10進表記	10進表記															
0	1	2	4	8	16	32	64	128	256	512	1024	2048					
1	14	28	56	112	224	448	896	1792	3584	3073	2051	7					
2	50	100	200	400	800	1600	3200	2305	515	1030	2060	25					
3	61	122	244	488	976	1952	3904	3713	3331	2567	1039	2078					
4	84	168	336	672	1344	2688	1281	2562	1029	2058	21	42					
:	:																
127	4094	4093	4091	4087	4079	4063	4031	3967	3839	3583	3071	2047					

【 図 1 6 】



【 図 1 7 】



【図 18】

符号変換テーブル

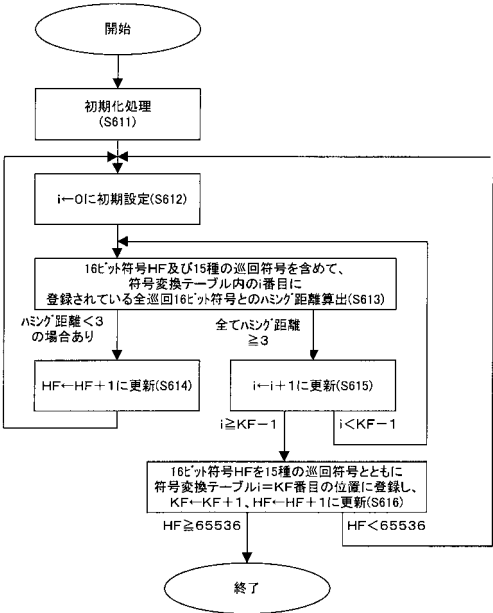
7ビット符号	16ビット符号	
10進表記	10進表記	2進表記
0	1	0000 0000 0000 0001
1	62	0000 0000 0011 1110
2	454	0000 0001 1100 0110
3	505	0000 0001 1111 1001
4	1610	0000 0110 0100 1010
5	1653	0000 0110 0111 0101
⋮	⋮	⋮
124	65107	1111 1110 0101 0011
125	65132	1111 1110 0110 1100
126	65428	1111 1111 1001 0100
127	65451	1111 1111 1010 1011

【図 19】

巡回符号表

7ビット符号 (基準符号)	16ビット符号 (ハミング符号)	巡回ビットパターン(15種類)
10進表記	10進表記	2進表記
0	1	2 4 8 16 32 64 128 256 512 1024 2048 4096 8192 16384 32768
1	62	124 248 496 992 1984 3968 7936 15872 31744 63488 61441 57347 49159 32783 31
2	454	908 1816 3632 7264 14528 29056 58112 50689 35843 6151 12302 24604 49208 32881 227
3	505	1010 2020 4040 8080 16160 32320 64640 63745 61955 58375 51215 36895 8255 16510 33020
4	1610	3220 6440 12880 25760 51520 37505 9475 18950 37900 10265 20530 41060 16585 33170 805
⋮	⋮	⋮
127	65451	65367 65199 64863 64191 62847 60159 54783 44031 22527 45054 24573 49146 32757 65514 65493

【図 20】



【図 21】

符号変換テーブル

8ビット符号	16ビット符号	
10進表記	10進表記	2進表記
0	1	0000 0000 0000 0001
1	15	0000 0000 0000 1111
2	51	0000 0000 0011 0011
3	85	0000 0000 0101 0101
4	91	0000 0000 0101 1011
5	157	0000 0000 1001 1101
⋮	⋮	⋮
129	22363	0101 0111 0101 1011
130	23295	0101 1010 1111 1111
131	28655	0110 1111 1110 1111
132	65535	1111 1111 1111 1111

【図 22】

巡回符号表

8ビット符号 (基準符号)	16ビット符号 (ハミング符号)	巡回ビットパターン(15種類)
10進表記	10進表記	2進表記
0	1	2 4 8 16 32 64 128 256 512 1024 2048 4096 8192 16384 32768
1	15	30 60 120 240 480 960 1920 3840 7680 15360 30720 61440 57345 49155 32775
2	51	102 204 408 816 1632 3264 6528 13056 26112 52224 38913 12291 24582 49164 32793
3	85	170 340 680 1360 2720 5440 10880 21760 43520 21505 43010 20485 40970 16405 32810
4	91	182 364 728 1456 2912 5824 11648 23296 46592 27649 55298 45061 24587 49174 32813
:	:	:
131	28655	57310 49085 32635 65270 65005 64475 63415 61295 57055 48575 31615 63230 60925 56315 47095
132	65535	65535