(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2012-141631
(P2012-141631A)
(43)公開日　平成24年7月26日(2012.7.26)

| (51)Int.Cl. | | F I | | テーマコード（参考） |
|---|---|---|---|---|
| G10L 21/04 | (2006.01) | G10L | 21/04 | 200C |
| G10L 11/00 | (2006.01) | G10L | 21/04 | 110Z |
| | | G10L | 21/04 | 120D |
| | | G10L | 11/00 | 403 |

審査請求 有　請求項の数 9　OL　外国語出願　（全 25 頁）

(21)出願番号　特願2012-55130 (P2012-55130)
(22)出願日　平成24年3月12日 (2012.3.12)
(62)分割の表示　特願2010-550054 (P2010-550054)の分割
原出願日　平成21年2月17日 (2009.2.17)
(31)優先権主張番号　61/035,317
(32)優先日　平成20年3月10日 (2008.3.10)
(33)優先権主張国　米国 (US)

(71)出願人　591037214
フラウンホッファー－ゲゼルシャフト　ツァ　フェルダールング　デァ　アンゲヴァンテン　フォアシュンク　エー．ファオ
ドイツ連邦共和国　８０６８６　ミュンヘン　ハンザシュトラッセ　２７ツェー
(74)代理人　100079577
弁理士　岡田　全啓
(72)発明者　ディスヒ　サッシャ
ドイツ連邦共和国　９０７６３　フルストゥルンシュトラーセ　７
(72)発明者　ナーゲル　フレデリック
ドイツ連邦共和国　９０４２５　ニュルンベルグ　ヴィルヘルムスハーフェナー　シュトラーセ　７２

最終頁に続く

(54)【発明の名称】瞬間的事象を有する音声信号の操作装置および操作方法

(57)【要約】　（修正有）
【課題】瞬間的事象を含む信号に音声効果を適用する。
【解決手段】瞬間的事象を有する音声信号を操作するための信号操作器は、瞬間的事象信号除去器１００、信号処理器１１０、および信号挿入器１２０を含む。信号操作器は、瞬間的事象信号除去器１００によって瞬間的事象が取り除かれた信号位置にて、処理音声信号に時間部分を挿入する。その結果、操作された音声信号は、処理によって影響を受けなかった瞬間的事象を含む。ここで、瞬間的事象の垂直コヒーレンスは、信号処理器１１０の中で実行される処理が、垂直コヒーレンスを破壊しないで維持される。
【選択図】図１

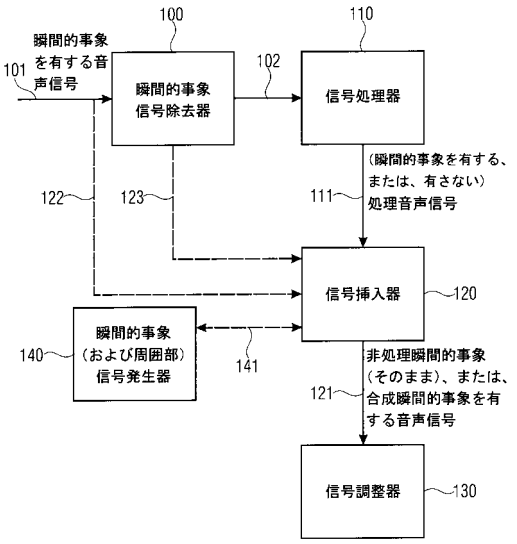



FIGURE 1

【特許請求の範囲】
【請求項１】
瞬間的事象（８０１）を有する音声信号の操作装置であって、
処理音声信号を得るために、前記瞬間的事象（８０１）を含む第１の時間部分（８０４）が除去された瞬間的事象減少の音声信号を処理するための、もしくは、前記瞬間的事象（８０１）を含む音声信号を処理するための、信号処理器（１１０）と、
第２の時間部分（８０９）を、前記第１の時間部分（８０４）が除去された、もしくは、前記瞬間的事象（８０１）が前記処理音声信号の中に置かれた信号位置で、前記処理音声信号に挿入するための信号挿入器（１２０）と、を備え、
前記瞬間的事象（８０１）を含む第２の時間部分（８０９）は、前記信号処理器（１１０）によって実行される処理によって影響をされず、その結果、操作音声信号が得られること、を特徴とし、
音声信号に関連したサイド情報を抽出して解読するためのサイド情報抽出器（１０６）を更に備え、
前記サイド情報は、前記瞬間的事象（８０１）の時間位置（８０３）を示す、または、前記第１の時間部分（８０４）もしくは前記第２の時間部分（８０９）の開始時間もしくは終了時間を示すこと、
を特徴とする、音声信号の操作装置。
【請求項２】
前記瞬間的事象減少の音声信号を得るために、前記音声信号から前記第１の時間部分（８０４）を除去するための瞬間的事象信号除去器（１００）を更に備え、前記第１の時間部分（８０４）は前記瞬間的事象（８０１）を有していること、を特徴とする、請求項１に記載の音声信号の操作装置。
【請求項３】
前記信号処理器（１１０）は、前記瞬間的事象減少の音声信号を、周波数依存方法（１１２、１１３）で処理するように構成され、その結果、前記処理は、前記瞬間的事象減少の音声信号に位相シフトを導入し、前記位相シフトは異なるスペクトル成分ごとに異なること、を特徴とする、請求項１または請求項２に記載の音声信号の操作装置。
【請求項４】
前記信号挿入器（１２０）は、少なくとも前記第１の時間部分（８０４）を複製することによって、前記第２の時間部分（８０９）を発生させるように構成され、その結果、前記第２の時間部分（８０９）は、前記瞬間的事象（８０１）を有する音声信号から複製した少なくとも第１の時間部分（８０４）の複製を含むこと、を特徴とする、請求項１～請求項３のいずれかに記載の音声信号の操作装置。
【請求項５】
前記信号処理器（１１０）は、音声分析合成装置、位相音声分析合成装置、もしくは、（Ｐ）ＳＯＬＡ処理器を含むこと、を特徴とする、請求項１～請求項４のいずれかに記載の音声信号の操作装置。
【請求項６】
時間が離散した前記操作音声信号の減衰もしくは挿入によって前記操作音声信号を調整するための信号調整器（１３０）を更に備えたこと、を特徴とする、請求項１～請求項５のいずれかに記載の音声信号の操作装置。
【請求項７】
音声信号における瞬間的事象を検出するための瞬間的事象検出器（１０３）、もしくは、音声信号に関連したサイド情報を抽出して解読するためのサイド情報抽出器（１０６）を更に備え、
前記サイド情報は、前記瞬間的事象（８０１）の時間位置（８０３）を示す、または、前記第１の時間部分（８０４）もしくは前記第２の時間部分（８０９）の開始時間もしくは終了時間を示すこと、
を特徴とする、請求項１～請求項６のいずれかに記載の音声信号の操作装置。

【請求項８】
瞬間的事象（８０１）を有する音声信号の操作方法であって、
処理音声信号を得るために、前記瞬間的事象（８０１）を含む第１の時間部分（８０４）が除去された瞬間的事象減少の音声信号を処理するための、もしくは、前記瞬間的事象（８０１）を含む音声信号を処理するための、信号処理工程（１１０）と、
第２の時間部分（８０９）を、前記第１の時間部分（８０４）が除去された、もしくは、前記瞬間的事象（８０１）が前記処理音声信号の中に置かれた信号位置で、前記処理音声信号に挿入するための信号挿入工程（１２０）と、を備え、
前記瞬間的事象（８０１）を含む第２の時間部分（８０９）は前記信号処理工程（１１０）によって影響されず、その結果、操作音声信号が得られること、を特徴とし、
音声信号に関連したサイド情報を抽出して解読する工程（１０６）を更に備え、
前記サイド情報は、前記瞬間的事象（８０１）の時間位置（８０３）を示す、または、前記第１の時間部分（８０４）もしくは前記第２の時間部分（８０９）の開始時間もしくは終了時間を示すこと、
を特徴とする、音声信号の操作方法。
【請求項９】
コンピュータが稼動したとき、請求項８の方法を実行するためのプログラムコードを有すること、を特徴とする、コンピュータプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、音声信号処理に関し、特に、瞬間的事象を含む信号に音声効果を適用する状況下での音声信号操作に関する。
【背景技術】
【０００２】
音声信号を操作して、ピッチを維持しながら再生速度が変えられることが知られている。そのような手順に関する周知の方法は、例えば、Ｊ．Ｌ．フラナガン（およびＲ．Ｍ．ゴールデンら著、ベルシステム技術ジャーナル、１９６６年１１月、ｐｐ１３９４～１５０９、米国特許第６５４９８８４号公報　ラロッシュ．Ｊ（Ｌａｒｏｃｈｅ．Ｊ）およびドルセン．Ｍ（Ｄｏｌｓｏｎ．Ｍ）「位相音声分析合成装置のピッチシフト」、ジーン・ラロッシュおよびマーク・ドルセン、「ピッチシフトのための新しい位相音声分析合成装置のテクニック、調和、および他のエキゾチックな効果」、音声と音響の信号処理の応用に関する１９９９年ＩＥＥＥ研究集の会報、ニュープラッツ、ニューヨーク１９９９年１０月１７日～２０日、ゼルザー．Ｕ著：ＤＡＦＸ：デジタル音声効果、ワイリーと息子、第１版、２００２年２月２６日、ページ２０１～２９８で説明されるように、位相音声分析合成装置、または、（ピッチ同期）重複加算法（（ｐｉｔｃｈ　ｓｙｎｃｈｒｏｎｏｕｓ）　ｏｖｅｒｌａｐ－ａｄｄ法、略して（Ｐ）ＳＯＬＡ法）のような方法によって実行される。
【０００３】
さらに、音声信号は、そのような方法、すなわち、位相音声分析合成装置または（Ｐ）ＳＯＬＡ法を使用して、転移させることができる。この種の転移の特に注目すべき点は、転移した音声信号は、ピッチは変更されているけれども、転移の前の元の音声信号と同じ再現／再生の長さを有しているということである。このことは、拡張された音声信号を、加速して再現することによって得られる。ここで、加速して再現することを実行するための加速係数は、時間において、元の音声信号を拡張するための拡張係数に依存する。転移した音声信号が、時間が離散した信号表現を有するとき、この手順は、サンプリング周波数が維持される拡張係数と等しい係数によって、拡張された音声信号の低標本抽出、または、拡張された音声信号の減衰に対応する。
【０００４】
そのような音声信号操作における特別の挑戦が、瞬間的事象である。瞬間的事象は、全

部の帯域、または、所定の周波数領域の信号エネルギーが、急激に変化する、すなわち、急激に増加または減少する、信号中の事象である。特別な瞬間的事象の特徴は、スペクトルおける信号エネルギーの分配である。通常、瞬間的事象期間中の音声信号のエネルギーは、周波数全体にわたって分配される。一方、非瞬間事象の信号部分において、信号エネルギーは、通常、音声信号の低周波数部分または特定の帯域に集中する。これは、静止信号部分または色調信号部分と称される非瞬間的事象信号部分が、平坦でないスペクトルを有することを意味する。言い換えれば、信号エネルギーは、音声信号の雑音床にわたって強く立ち上がる、比較的小さい数のスペクトル線／スペクトル帯域に含まれる。しかしながら、瞬間的事象部分の中では、音声信号のエネルギーが、多くの異なる周波数帯域に、特に、高周波部分に分配される。その結果、音声信号の瞬間的事象部分のスペクトルは比較的平坦であり、音声信号の色調部分のスペクトルより、とにかく平坦である。通常、瞬間的事象は、時間内に激しく変化する。それは、フーリエ分解が実行されるとき、信号が多くの高調波を含むことを意味する。これらの多くの高調波の重要な特徴は、これらの高調波の位相が非常に特別な相互関係にあるということである。その結果、これらすべての正弦波の重ね合わせが、信号エネルギーの急激な変化をもたらす。言い換えれば、スペクトル相互に強い相関関係が存在する。

【０００５】

また、すべての階調波の間の特別な位相状況は、「垂直コヒーレンス（ｖｅｒｔｉｃａｌ　ｃｏｈｅｒｅｎｃｅ）」と称することができる。この「垂直コヒーレンス」は、信号の時間／周波数スペクトル表示に関するものであり、周波数の短時間スペクトルにおいて、横軸方向が時間における信号の進展に対応し、垂直軸の寸法がスペクトル成分（変換周波数ビン（ｂｉｎ））の周波数における相互依存を示す。

【０００６】

音声信号の時間を拡張または縮小するために実行される通常の処理ステップにより、この垂直コヒーレンスが破壊される。これは、瞬間的事象が、時間の拡張操作または短縮操作されるとき、瞬間的事象が、時間経過により「塗り付けられる」ことを意味する。時間の拡張操作または短縮操作は、例えば、位相音声分析合成装置または別の方法によって実行される。位相音声分析合成装置または別の方法は、音声信号に位相シフトを導入する周波数依存処理を実行する。位相シフトは、異なる周波数係数ごとに異なる。

【０００７】

瞬間的事象の垂直コヒーレンスが、音声信号を処理する方法によって破壊されるとき、操作された信号は、静止部分または非瞬間的事象部分において、元の信号と非常に似たものとなる。しかし、操作された信号において、瞬間的事象部分は品質が低下する。瞬間的事象の垂直コヒーレンスの非制御の操作は、瞬間的事象の一時的な分散をもたらす。多くの高調波の成分が瞬間的事象に貢献し、非制御の方法でこれらのすべての高調波の成分の位相を変更することは、このような人工物（分散）を必然的にもたらす。

【０００８】

しかしながら、瞬間的事象部分は、音楽信号やスピーチ信号のような、動的な音声信号にとって非常に重要である。特定の時間内の音声エネルギーの突然の変化は、操作された信号の品質において非常に多くの主観的なユーザの印象を表す。言い換えれば、音声信号における瞬間的事象は、通常、音声信号のかなり顕著な「重大事件」であり、主観的な品質の印象に過剰に比例した影響を与える。垂直コヒーレンスが、信号処理操作によって破壊され、または、元の信号の瞬間的事象部分に関して低下した、操作された瞬間的事象は、聴衆にとって、歪んで、反響して、そして不自然に聞こえる。

【０００９】

いくつかの現行手法は、瞬間的事象の期間の間、時間拡張が無い、または、時間拡張がより少ない実行を継続してしなければならないように、瞬間的事象の周囲の時間を、より高い程度まで拡張する。そのような従来技術の文献および特許が、時間、および／または、ピッチ操作の方法を説明する。従来技術の文献は、ラロッシュ．Ｌおよびドルセン．Ｍ、「音声の改良された位相音声分析合成装置の時間スケール変更」、ＩＥＥＥ通信、スピ

ーチおよび音声処理、７巻、Ｎｏ．３、ページ３２３～３３２、エマニュエル・ラベリ、マーク・サンドラーおよびホアン・Ｐ．ベロ、ステレオ音声の非線形の時間スケールの高速実行、デジタル音声効果の第８回国際会議（ＤＡＦｘ´０５）の議事録、マドリード、スペイン、２００５年９月２０日～２２日、ダックスブリ、Ｃ．Ｍ．デイヴィースおよびＭ．サンドラー（２００１年、１２月）、マルチ解決分析技術を使用した、音楽音声の瞬間的事象情報の分離、デジタル音声効果のＣＯＳＴ　Ｇ－６会議（ＤＡＦＸ－０１）の議事録、リムリック、アイルランド、およびローベル、Ａ．：位相音声分析合成装置での瞬間的事象の処理に対する新しいアプローチ、デジタル音声効果の第６回国際会議（ＤＡＦｘ－０３）の議事録、ロンドン、イギリス、２００３年９月８日～１１日である。

【先行技術文献】

【特許文献】

【００１０】

【特許文献１】米国特許第６５４９８８４号

【非特許文献】

【００１１】

【非特許文献１】Ｊ．Ｌ．フラナガンおよびＲ．Ｍ．ゴールデン、ベルシステム技術ジャーナル、１９６６年１１月、ページ１３９４～１５０９（Ｊ．Ｌ．Ｆｌａｎａｇａｎ ａｎｄ Ｒ．Ｍ．Ｇｏｌｄｅｎ，Ｔｈｅ Ｂｅｌｌ Ｓｙｓｔｅｍ Ｔｅｃｈｎｉｃａｌ Ｊｏｕｒｎａｌ，Ｎｏｖｅｍｂｅｒ １９６６，ｐｐ．１３９４ ｔｏ １５０９）

【非特許文献２】ジーン・ラロッシュおよびマーク・ドルセン、「ピッチシフトのための新しい位相音声分析合成装置のテクニック、調和、および他のエキゾチックな効果」、音声と音響の信号処理の応用に関する１９９９年ＩＥＥＥ研究集の会報、ニュープラッツ、ニューヨーク１９９９年１０月１７日～２０日（Ｊｅａｎ Ｌａｒｏｃｈｅ ａｎｄ Ｍａｒｋ Ｄｏｌｓｏｎ，Ｎｅｗ Ｐｈａｓｅ－Ｖｏｃｏｄｅｒ Ｔｅｃｈｎｉｑｕｅｓ ｆｏｒ Ｐｉｔｃｈ－Ｓｈｉｆｔｉｎｇ，Ｈａｒｍｏｎｉｚｉｎｇ Ａｎｄ Ｏｔｈｅｒ Ｅｘｏｔｉｃ Ｅｆｆｅｃｔｓ”，Ｐｒｏｃ．１９９９ ＩＥＥＥ Ｗｏｒｋｓｈｏｐ ｏｎ Ａｐｐｌｉｃａｔｉｏｎｓ ｏｆ Ｓｉｇｎａｌ Ｐｒｏｃｅｓｓｉｎｇ ｔｏ Ａｕｄｉｏ ａｎｄ Ａｃｏｕｓｔｉｃｓ，Ｎｅｗ Ｐａｌｔｚ，Ｎｅｗ Ｙｏｒｋ，Ｏｃｔ．１７－２０，１９９９）

【非特許文献３】ゼルザー．Ｕ著：ＤＡＦＸ：デジタル音声効果、ワイリーと息子、第１版、２００２年２月２６日、ページ２０１～２９８（Ｚｏｌｚｅｒ，Ｕ：ＤＡＦＸ：Ｄｉｇｉｔａｌ Ａｕｄｉｏ Ｅｆｆｅｃｔｓ；Ｗｉｌｅｙ ＆ Ｓｏｎｓ；Ｅｄｉｔｉｏｎ：１（Ｆｅｂｒｕａｒｙ ２６，２００２）； ｐｐ． ２０１－２９８）

【非特許文献４】ラロッシュ．Ｌおよびドルセン．Ｍ、「音声の改良された位相音声分析合成装置の時間スケール変更」、ＩＥＥＥ通信、スピーチおよび音声処理、７巻、Ｎｏ．３、ページ３２３～３３２（Ｌａｒｏｃｈｅ Ｌ．，Ｄｏｌｓｏｎ Ｍ．：Ｉｍｐｒｏｖｅｄ ｐｈａｓｅ ｖｏｃｏｄｅｒ ｔｉｍｅｓｃａｌｅ ｍｏｄｉｆｉｃａｔｉｏｎ ｏｆ ａｕｄｉｏ”，ＩＥＥＥ Ｔｒａｎｓ． Ｓｐｅｅｃｈ ａｎｄ Ａｕｄｉｏ Ｐｒｏｃｅｓｓｉｎｇ，ｖｏｌ．７，ｎｏ．３，ｐｐ．３２３－３３２）

【非特許文献５】エマニュエル・ラベリ、マーク・サンドラーおよびホアン・Ｐ．ベロ、ステレオ音声の非線形の時間スケールの高速実行、デジタル音声効果の第８回国際会議（ＤＡＦｘ´０５）の議事録、マドリード、スペイン、２００５年９月２０日～２２日（Ｅｍｍａｎｕｅｌ Ｒａｖｅｌｌｉ，Ｍａｒｋ Ｓａｎｄｌｅｒ ａｎｄ Ｊｕａｎ Ｐ．Ｂｅｌｌｏ：Ｆａｓｔ ｉｍｐｌｅｍｅｎｔａｔｉｏｎ ｆｏｒ ｎｏｎ－ｌｉｎｅａｒ ｔｉｍｅ－ｓｃａｌｉｎｇ ｏｆ ｓｔｅｒｅｏ ａｕｄｉｏ；Ｐｒｏｃ．ｏｆ ｔｈｅ ８ｔｈ Ｉｎｔ．Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ｄｉｇｉｔａｌ Ａｕｄｉｏ Ｅｆｆｅｃｔｓ （ＤＡＦｘ’０５），Ｍａｄｒｉｄ，Ｓｐａｉｎ，Ｓｅｐｔｅｍｂｅｒ ２０－２２，２００５）

【非特許文献６】ダックスブリ、Ｃ．Ｍ．デイヴィースおよびＭ．サンドラー（２００１年、１２月）、マルチ解決分析技術を使用した、音楽音声の瞬間的事象情報の分離、デジ

タル音声効果のＣＯＳＴ　Ｇ－６会議（ＤＡＦＸ－０１）の議事録、リムリック、アイルランド（Ｄｕｘｂｕｒｙ，　Ｃ．Ｍ．Ｄａｖｉｅｓ，ａｎｄ Ｍ．Ｓａｎｄｌｅｒ（２００１，Ｄｅｃｅｍｂｅｒ）．Ｓｅｐａｒａｔｉｏｎ ｏｆ ｔｒａｎｓｉｅｎｔ ｉｎｆｏｒｍａｔｉｏｎ ｉｎ ｍｕｓｉｃａｌ ａｕｄｉｏ ｕｓｉｎｇ ｍｕｌｔｉｒｅｓｏｌｕｔｉｏｎ ａｎａｌｙｓｉｓ ｔｅｃｈｎｉｑｕｅｓ．Ｉｎ Ｐｒｏｃｅｅｄｉｎｇｓ ｏｆ ｔｈｅ ＣＯＳＴ Ｇ－６ Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ｄｉｇｉｔａｌ Ａｕｄｉｏ Ｅｆｆｅｃｔｓ（ＤＡＦＸ－０１），Ｌｉｍｅｒｉｃｋ，Ｉｒｅｌａｎｄ）

【非特許文献７】ローベル、Ａ．：位相音声分析合成装置での瞬間的事象の処理に対する新しいアプローチ、デジタル音声効果の第６回国際会議（ＤＡＦｘ－０３）の議事録、ロンドン、イギリス、２００３年9月8日～１１日（Ｒｏｂｅｌ，　Ａ．：Ａ ＮＥＷ ＡＰＰＲＯＡＣＨ ＴＯ ＴＲＡＮＳＩＥＮＴ ＰＲＯＣＥＳＳＩＮＧ ＩＮ ＴＨＥ ＰＨＡＳＥ ＶＯＣＯＤＥＲ；Ｐｒｏｃ．ｏｆ ｔｈｅ ６ｔｈ Ｉｎｔ．Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ｄｉｇｉｔａｌ Ａｕｄｉｏ Ｅｆｆｅｃｔｓ （ＤＡＦｘ－０３），Ｌｏｎｄｏｎ，ＵＫ，Ｓｅｐｔｅｍｂｅｒ ８－１１，２００３）

【発明の概要】

【発明が解決しようとする課題】

【００１２】

位相音声分析合成装置による音声信号の時間拡張の間、瞬間的事象部分は、分散によって「ぼかされる」。いわゆる信号の垂直コヒーレンスが損なわれるからである。（Ｐ）ＳＯＬＡ法のような、いわゆる重複加算方法を使用する方法は、瞬間的事象の前エコーおよび後エコーの擾乱を発生させる。これらの問題は、瞬間的事象の周囲で時間拡張を増加させることによって、実際に記述される。しかしながら、仮に、転移が起こるならば、転移係数は、もはや瞬間的事象の周囲で一定にならない。すなわち、重畳された（色調の）信号成分のピッチは変化して、擾乱として知覚される。

【００１３】

それゆえに、本発明の主たる目的は、より高い品質の操作された音声信号が得られる、瞬間的事象を有する音声信号の操作装置および操作方法を提供することである。

【課題を解決するための手段】

【００１４】

この目的は、請求項１に記載の音声信号の操作装置、請求項８に記載の音声信号の操作方法および請求項９に記載のコンピュータプログラムによって達成される。

【００１５】

瞬間的事象部分の非制御の処理の中で起こる品質問題を処理するために、本発明は、瞬間的事象部分が有害な方法で全く処理されないということを確実にする。すなわち、瞬間的事象は、処理の前に除去され、処理した後に再び挿入される、あるいは、瞬間的事象は、処理されるが、処理音声信号から除去され、非処理の瞬間的事象に置き換えられる。

【００１６】

好ましくは、処理音声信号に挿入された瞬間的事象部分は、元の音声信号の対応する瞬間的事象部分の複製である。その結果、操作音声信号は、瞬間的事象を含まない処理部分と、瞬間的事象を含む非処理部分または異処理部分と、で構成される。例えば、元の瞬間的事象は、減衰、ある種の重み付け、または、パラメータ化処理がされる。しかしながら、瞬間的事象部分は、合成して作成された瞬間的事象部分に置き換えられる。合成された瞬間的事象部分は、所定の時間内において変化するエネルギー量などのいくつかの瞬間的事象パラメータ、または、瞬間的事象を特徴付ける別の測度に関して、合成された瞬間的事象部分が元の瞬間的事象部分と同様であるような方法で合成される。その結果、１つには、元の音声信号の瞬間的事象部分を特徴付けでき、また、１つには、この瞬間的事象を処理の前に除去したり、処理された瞬間的事象を合成された瞬間的事象に置き換えたりできる。合成された瞬間的事象は、瞬間的事象パラメータ情報に基づいて、合成的に作成される。しかしながら、効率の理由で、操作の前に元の音声信号の一部を複製して、この複

製を処理音声信号に挿入することが好ましい。この手順は、処理音声信号の瞬間的事象部分が、元の音声信号の瞬間的事象と同じであることを保証するからである。この手順は、処理前の元の音声信号と比較される処理音声信号において、音響信号知覚の瞬間的事象の特別に高い影響が維持されることを確実なものとする。したがって、瞬間的事象に関する主観的または客観的な品質は、音声信号を操作するための、ある種の音声信号処理によって低下しない。

【００１７】

好ましい実施形態において、本発明は、そのような処理の枠組みの中で、瞬間的音声事象の知覚の優遇のための新しい方法を提供する。そうでなければ、枠組みは、音声信号の分散によって一時的な「手ぶれ」を発生させる。この好ましい方法は、時間拡張の目的のために、信号操作の前に瞬間的音声事象の除去を本質的に含み、次に、拡張を考慮に入れながら、非処理の瞬間的事象信号部分を、正確な方法で変更された（拡張された）音声信号に加える。

【００１８】

以下に、本発明の好適な実施形態が添付図面を参照してより詳細に説明される。

【図面の簡単な説明】

【００１９】

【図１】図１は、瞬間的事象を有する音声信号の操作装置または操作方法の好ましい実施形態を示すブロック図である。

【図２】図２は、図１の瞬間的事象信号除去器の好ましい実施例を示すブロック図である。

【図３Ａ】図３Ａは、図１の信号処理器の好ましい実施例を示すブロック図である。

【図３Ｂ】図３Ｂは、図１の信号処理器の別の好ましい実施例を示すブロック図である。

【図４】図４は、図１の信号挿入器の好ましい実施例を示すブロック図である。

【図５Ａ】図５Ａは、図１の信号処理器で使用される音声分析合成装置の概略実施例を示すブロック図である。

【図５Ｂ】図５Ｂは、図１の信号処理器の一部（分析部）の実施例を示すブロック図である。

【図５Ｃ】図５Ｃは、図１の信号処理器の別の一部（拡張部）の実施例を示すブロック図である。

【図５Ｄ】図５Ｄは、図１の信号処理器の別の一部（合成部）の実施例を示すブロック図である。

【図６】図６は、図１の信号処理器で使用される位相音声分析合成装置の変換実施例を示すブロック図である。

【図７Ａ】図７Ａは、帯域幅拡大処理構成の符号器側を示すブロック図である。

【図７Ｂ】図７Ｂは、帯域幅拡大処理構成の復号器側を示すブロック図である。

【図８Ａ】図８Ａは、瞬間的事象を伴う音声入力信号のエネルギー表示を示すグラフである。

【図８Ｂ】図８Ｂは、窓のある瞬間的事象を伴う、図８Ａの音声入力信号のエネルギー表示を示すグラフである。

【図８Ｃ】図８Ｃは、拡張される前の、瞬間的事象部分の無い音声入力信号のエネルギー表示を示すグラフである。

【図８Ｄ】図８Ｄは、拡張された後の、図８Ｃの音声入力信号のエネルギー表示を示すグラフである。

【図８Ｅ】図８Ｅは、元の音声入力信号の対応部分が挿入された後の、操作音声入力信号のエネルギー表示を示すグラフである。

【図９】図９は、音声信号のためのサイド情報発生装置を示すブロック図である。

【発明を実施するための形態】

【００２０】

図１は、瞬間的事象を有する音声信号を操作するための好ましい装置を示す。この装置

は、瞬間的事象を伴う音声信号の入力１０１を有する、瞬間的事象信号除去器１００を含む。瞬間的事象信号除去器１００の出力１０２は、信号処理器１１０に接続されている。信号処理器１１０の出力１１１は、信号挿入器１２０に接続されている。信号挿入器１２０の出力１２１では、非処理（「そのまま」）または合成された瞬間的事象を伴う操作音声信号が得られ、信号調整器１３０などの別の装置に接続される。信号調整器１３０は、図７ａおよび図７ｂに関係して議論するように、帯域幅拡大目的に必要である低標本抽出／減衰などの、操作音声信号の更なる処理を実行できる。

【００２１】

しかしながら、仮に、信号挿入器１２０の出力１２１で得られる操作音声信号が、そのまま使用され、すなわち、更なる処理のために格納され、または、受信器に送られ、または、デジタル／アナログ変換器に送られ、最後は、スピーカ設備に接続され、操作音声信号を表す音響信号を最終的に発生させるならば、信号調整器１３０は全く使用されない。

【００２２】

帯域幅拡大の場合において、信号線（出力）１２１の音声信号は、既に高帯域信号である。信号処理器１１０は、入力低帯域信号から高帯域信号を発生させる。そして、音声信号の入力１０１から抽出された低帯域の瞬間的事象部分は、高帯域の周波数領域の中に置かなければならない。それは、好ましくは、減衰のような、垂直コヒーレンスを妨げない信号処理によって行われる。減衰は信号挿入器１２０の前で実行され、減衰された瞬間的事象部分は、信号処理器１１０の出力１１１で、高帯域信号に挿入される。本実施形態では、信号調整器１３０は、例えば、ＭＰＥＧ４のスペクトル帯域複製（Ｓｐｅｃｔｒａｌ　Ｂａｎｄ　Ｒｅｐｌｉｃａｔｉｏｎ）の中で行われるような、エンベロープ形成、雑音加算、階調波の逆フィルタリングや加算などの高帯域信号の別の処理を実行する。

【００２３】

信号挿入器１２０は、好ましくは、信号線１２３を通して瞬間的事象信号除去器１００からサイド情報を受信し、非処理の信号から正しい部分を選んで、出力１１１に挿入する。

【００２４】

装置１００，１１０，１２０，１３０を有する実施形態が実行されるとき、図８ａ～図８ｅに関して議論される信号シーケンスが得られる。しかしながら、信号処理器１１０で信号処理操作を実行する前に、瞬間的事象部分を除去することは、必ずしも必要ではない。本実施形態において、瞬間的事象信号除去器１００は必要でなく、信号挿入器１２０は、出力１１１の処理音声信号から切り外すべき信号部分を決定し、この切り外した信号部分を、信号線１２１によって図式的に示される元の信号の一部で置き換えること、あるいは、信号線１４１によって示される合成信号で置き換えることを決定する。この合成信号は瞬間的事象信号発生器１４０で発生される。適した瞬間的事象を発生できるように、信号挿入器１２０は、瞬間的事象記述パラメータを瞬間的事象信号発生器１４０に伝達するように構成されている。したがって、矢印１４１によって示される瞬間的事象信号発生器１４０と信号挿入器１２０との間の接続は、双方向接続として記載されている。特定の瞬間的事象検出器１０３（図１では、図示しない）が、操作装置の中に設けられているときは、瞬間的事象部分の情報が、この瞬間的事象検出器１０３から瞬間的事象信号発生器１４０に提供される。瞬間的事象信号発生器１４０は、直接使用できる瞬間的事象のサンプルを有したり、あるいは、予め格納された瞬間的事象のサンプルを有したりするように構成してもよい。瞬間的事象のサンプルは、信号挿入器１２０によって使用される瞬間的事象を実際に発生／合成するために、瞬間的事象パラメータを使用して重み付けされる。

【００２５】

本実施形態において、瞬間的事象信号除去器１００は、音声信号から第１の時間部分を除去して、瞬間的事象部分が減少した音声信号を得るように構成されている。ここで、第１の時間部分は瞬間的事象を含む。

【００２６】

さらに、好ましくは、信号処理器１１０は、出力１１１の処理音声信号を得るために、

瞬間的事象を含む第１の時間部分が除去された瞬間的事象減少の音声信号を処理するように、または、瞬間的事象を含む音声信号を処理するように、構成される。

【００２７】

好ましくは、信号挿入器１２０は、第１の時間部分が除去された信号位置で、または、瞬間的事象が音声信号の中で位置している信号位置で、第２の時間部分を、処理音声信号に挿入するように構成される。ここで、第２の時間部分は、信号処理器１１０によって実行された処理によって影響されない瞬間的事象を含む。その結果、出力１２１における操作音声信号が得られる。

【００２８】

図２は、瞬間的事象信号除去器１００の好ましい実施形態を示す。本実施形態において、音声信号は、瞬間的事象の少しのサイド情報／メタ情報も含んでいない。瞬間的事象信号除去器１００は、瞬間的事象検出器１０３、フェードアウト／フェードイン計算機１０４、第１の時間部分除去器１０５を含む。別の実施形態において、音声信号における瞬間的事象の情報が、後の図９に関係して議論する符号化装置によって、音声信号に付加されるように集められる。瞬間的事象信号除去器１００はサイド情報抽出器１０６を含む。サイド情報抽出器１０６は、信号線１０７によって示される音声信号に付加されたサイド情報を抽出する。瞬間的事象時間の情報は、信号線１０７によって示されるように、フェードアウト／フェードイン計算機１０４に提供される。しかしながら、音声信号が、メタ情報として、瞬間的事象の時間（すなわち、瞬間的事象が起こる正確な時間）（だけ）ではなく、音声信号から除かれる時間部分の開始／停止時間（すなわち、音声信号の「第１の時間部分」の開始時間と停止時間）を含むときは、フェードアウト／フェードイン計算機１０４は必要ではない。そして、開始／停止時間情報は、信号線１０８によって示されるように、直接に第１の時間部分除去器１０５に転送される。信号線１０８は任意であることを示す。また、破線によって示される他のすべての信号線も同様に、任意である。

【００２９】

図２において、フェードイン／フェードアウト計算機１０４は、好ましくは、サイド情報１０９を出力する。このサイド情報１０９は、第１の時間部分の開始／停止時間と異なる。図１の信号処理器１１０での処理の特性が、考慮されるからである。さらに、入力音声信号が、好ましくは、第１の時間部分除去器１０５に送られる。

【００３０】

好ましくは、フェードアウト／フェードイン計算機１０４は、第１の時間部分の開始／停止時間を提供する。これらの時間は、瞬間的事象の時間に基づいて計算される。その結果、瞬間的事象だけではなく、瞬間的事象の周囲のいくつかのサンプルも、第１の時間部分除去器１０５によって除去される。さらに、時間領域の矩形窓によって瞬間的事象部分を切り取らないで、フェードアウト部分およびフェードイン部分によって抽出を実行することが好ましい。フェードアウト部分またはフェードイン部分によって抽出を実行するためには、上昇する余弦波窓などの矩形のフィルタと比較して、滑らかな瞬間的事象を有する、どんな種類の窓も適用される。その結果、この抽出の周波数特性は、矩形窓が適用されたときほどの問題はない。なお、これは任意である。この時間領域の窓付け操作は、窓付け操作の残りの部分の音声信号、すなわち、窓が付けられた部分がない音声信号を出力する。

【００３１】

瞬間的事象の除去の後に、瞬間的事象減少の残留信号、または、好ましくは十分に瞬間的事象の無い残留信号を残す、そのような瞬間的事象の抑制方法が、この文脈の中において適用される。音声信号が時間の特定の部分にわたってゼロに設定される、瞬間的事象の完全な除去と比較して、瞬間的事象の抑制は、音声信号の更なる処理が、ゼロに設定された部分から損害を被る状況において有利である。そのようなゼロに設定された部分は、音声信号に対しては、非常に不自然だからである。

【００３２】

第１の時間部分の瞬間的事象の時間、および／または、開始／停止時間などのこれらの

計算の結果が、分離伝送チャンネルを通して伝達されるべき分離音声メタデータ信号などの、音声信号に伴って、または、音声信号とは別に、サイド情報またはメタ情報のいずれか一つとして、信号操作器に伝送される限り、当然、すべての計算が、図９に関係して議論する符号化側で同様に適用される、瞬間的事象検出器１０３およびフェードアウト／フェードイン計算機１０４によって実行される。

【００３３】

図３ａは、図１の信号処理器１１０の好ましい実施例を示す。この信号処理器１１０は、周波数選択分析器１１２と、次に接続された周波数選択処理装置１１３とを含む。周波数選択処理装置１１３は、元の音声信号の垂直コヒーレンスに負の影響を与えるように構成される。例えば、この処理は、音声信号の時間拡張、または、音声信号の時間短縮である。ここで、この拡張や短縮は、周波数選択方法で適用される。その結果、例えば、その処理は、処理音声信号に位相シフトを導入する。位相シフトは、異なる周波数帯域ごとに異なる。

【００３４】

処理の好ましい方法は、位相音声分析合成装置処理の文脈の中で、図３Ｂで示される。一般に、位相音声分析合成装置は、サブ帯域／変換分析器１１４と、サブ帯域／変換分析器１１４によって提供される複数の出力信号の周波数選択処理を実行するための、次に接続された処理器１１５と、後続のサブ帯域／変換合成器１１６と、を含む。サブ帯域／変換合成器１１６は、処理器１１５によって処理された信号を合成し、出力１１７で最終的に時間領域の処理信号を得る。ここで、出力１１７の処理信号の帯域幅が、処理器１１５とサブ帯域／変換合成器１１６との間の一つの分枝によって表された帯域幅より大きい限り、時間領域のこの処理信号は、再び完全な帯域幅信号または低帯域通過フィルタの信号である。サブ帯域／変換合成器１１６は、周波数選択信号の合成を実行する。

【００３５】

位相音声分析合成装置に関する詳細は、後で、図５Ａ、図５Ｂ、図５Ｃおよび図６に関連して議論する。

【００３６】

次に、図１の信号挿入器１２０の好ましい実施例が、図４で議論される。信号挿入器１２０は、好ましくは、第２の時間部分の長さを計算するための計算機１２２を含む。瞬間的事象が、図１の信号処理器１１０の中で信号処理される前に除去される実施形態において、第２の時間部分の長さが計算できるように、除去された第１の時間部分の長さと時間拡張係数（または、時間短縮係数）が必要である。その結果、第２の時間部分の長さが、計算機１２２の中で計算される。これらのデータ項目は、図１と図２で議論したように、外部から入力される。例示的に、第２の時間部分の長さは、第１の時間部分の長さを拡張係数に掛けることによって計算される。

【００３７】

第２の時間部分の長さは、音声信号における第２の時間部分の第１の境界と第２の境界とを計算するために、計算機１２３に伝送される。特に、計算機１２３は、入力１２４で供給される瞬間的事象の無い処理音声信号と、入力１２５で供給される瞬間的事象を伴う音声信号と、の間の相互相関処理を実行するように構成される。瞬間的事象を伴う音声信号は、第２の時間部分を提供する。好ましくは、計算機１２３は、別の制御入力１２６によって制御される。第２の時間部分の中の瞬間的事象の正シフトは、後で議論するように、瞬間的事象の負シフトに対して好ましい。

【００３８】

第２の時間部分の第１の境界と第２の境界は、抽出器１２７に提供される。好ましくは、抽出器１２７は、入力１２５で提供された元の音声信号から第２の時間部分を切り取る。その後、相互フェーダ１２８が使用されているので、切り取りは矩形のフィルタを使用して行われる。相互フェーダ１２８は、第２の時間部分の開始部分と第２の時間部分の停止部分とが、開始部分に対して０から１に増加する重み付けによって、および／または、終わりの部分に対して１から０に減少する重み付けによって、重み付けされる。その結果

、この相互フェード領域において、抽出信号の開始部分と共に処理信号の終わりの部分が加算されて、役に立つ信号をもたらす。同様の処理が、抽出後の処理音声信号の第２の時間部分の終わりの部分と始まりの部分とに対して、相互フェーダ１２８にて実行される。相互フェードは、瞬間的事象部分の無い処理音声信号の境界と第２の時間部分の境界とが完全に合致していないとき、クリックする人工物（分散）として、別の方法で知覚できる時間領域の人工物（分散）が発生しないことを確実にする。

【００３９】

次に、図５ａ、図５ｂ、図５ｃ、および図６を参照して、位相音声分析合成装置の文脈の中で、信号処理器１１０の好ましい実施例を説明する。

【００４０】

以下では、図５ａ、図５ｂ、図５ｃ、および図６を参照して、音声分析合成装置の好ましい実施例が、本発明に従って示される。図５ａは位相音声分析合成装置のフィルタバンクの実施例を示す。フィルタバンクにおいて、音声信号は、入力５００に送り込まれ、出力５１０にて得られる。特に、図５ａで示された概略的なフィルタバンクの各チャンネルは、帯域通過フィルタ５０１と下流の発振器５０２とを含む。すべてのチャンネルからのすべての発振器の出力信号は、合成器によって合成される。合成器は、出力信号を得るために、例えば、加算器として実行され、符号５０３で示される。各フィルタ５０１は、一方で振幅信号を、他方で周波数信号を供給するように構成される。振幅信号と周波数信号は時間信号である。振幅信号は、時間が経過するにつれてフィルタ５０１での振幅の進展を示す。一方、周波数信号は、フィルタ５０１によって篩にかけられた信号の周波数の進展を表す。

【００４１】

フィルタ５０１の概略的構成は、図５ｂで示される。図５ａの各フィルタ５０１は、図５ｂで示されるように構成される。しかしながら、そこでは、２つの入力混合器５５１および加算器５５２に供給した周波数ｆｉだけが、チャンネルごとに異なる。２つの入力混合器５５１の出力信号は、共に低帯域通過フィルタ５５３によって篩にかけられた低帯域通過信号である。２つの低帯域通過信号は、局部発振器周波数（ＬＯ周波数）によって発生する限り、位相が９０°異なる。上側の低帯域通過フィルタ５５３は直角位相信号５５４を提供し、一方、下側の低帯域通過フィルタ５５３は同相信号５５５を提供する。これらの２つの信号（すなわち、同相信号Ｉと直角位相信号Ｑ）は、矩形表現から大きさ位相表現を発生させる調整変換器５５６に提供される。時間が経過するにつれて、図５ａの大きさ信号または振幅信号が、それぞれ、出力５５７にて出力される。位相信号は、位相非包装器（ｐｈａｓｅ　ｕｎｗｒａｐｐｅｒ）５５８に提供される。位相非包装器５５８の出力において、直線的に増加する位相値の他には、常に０°～３６０°の間の現在の位相値はもはや存在しない。この「非包装」位相値は、位相／周波数変換器５５９に供給される。位相／周波数変換器５５９は、例えば、簡単な位相差形成器として構成され、現在の時点での位相から、種々の時点での位相を減算して、現在の時点の周波数値を得る。この周波数値は、フィルタチャンネルｉの一定の周波数値ｆｉに加算され、出力５６０にて一時的に変化する周波数値を得る。出力５６０における周波数値は、直接成分である平均周波数値（一定の周波数値）ｆｉと、選択成分であるフィルタチャンネルの信号の現在の周波数が平均周波数値ｆｉから外れた周波数偏差と、を有する。

【００４２】

したがって、図５ａと図５ｂで示されるように、位相音声分析合成装置はスペクトル情報と時間情報の分離を達成する。スペクトル情報は、特定のチャンネルの中に、または、周波数の直接成分を各チャンネルに供給する平均周波数値ｆｉの中に含まれる。一方、時間情報は、時間の経過に伴う周波数偏差または大きさの中に、それぞれ含まれる。

【００４３】

図５ｃは、本発明に従って、特に、音声分析合成装置の中で、かつ、図５ａの破線で示された回路の位置で、帯域幅増加を実行する操作を示す。

【００４４】

時間スケーリングに対して、例えば、各信号の中の信号ｆ（ｔ）のそれぞれのチャンネルまたは周波数の中の振幅信号Ａ（ｔ）は、減衰または挿入される。伝送の目的に対して、それが本発明の役に立つのであれば、挿入、すなわち、信号Ａ（ｔ）と信号ｆ（ｔ）の一時的な拡張または拡大が、拡張された信号Ａ’（ｔ）と信号ｆ’（ｔ）を得るために実行される。挿入が、帯域幅拡張のシナリオの中で拡張係数によって制御される。位相の変化の挿入、すなわち、加算器５５２による一定の周波数値ｆｉの加算の前の値によって、図５ａの個々の発振器５０２の周波数は変更されない。しかしながら、音声信号全体の一時的な変化は、すなわち、拡張係数２によって減速される。その結果は、元のピッチを有する一時的に拡張された音調、すなわち、階調波を伴う元の基本波である。

【００４５】

図５ｃで示された信号処理を実行することによって、そのような処理は、図５ａのすべてのフィルタ帯域チャンネルで実行される。決定器の中で決定された、結果である一時的な信号によって、音声信号は、すべての周波数が同時に２倍にされる期間中、元の信号に戻る。これは拡張係数２によるピッチ転移に導く。しかしながら、元の音声信号と同じ長さ、すなわち、同じ数のサンプルを有している音声信号が得られる。

【００４６】

また、図５ａで示されたフィルタバンクの実施例に代わるものとして、位相音声分析合成装置の変換構成が、図６に表現されるように使用される。ここで、音声信号１００は、一連の時間サンプルとして、ＦＦＴ処理器、または、より一般的に、短時間フーリエ変換処理器６００に供給される。ＦＦＴ処理器６００は図６の中に概略的に構成され、ＦＦＴによってスペクトルの大きさと位相を計算するために、音声信号の時間窓を実行する。この計算は、連続したスペクトルに対して実行される。連続したスペクトルは、強く重複している音声信号のブロックに関係する。

【００４７】

極端な場合は、あらゆる新しい音声信号のサンプルに対して、新しいスペクトルが計算される。新しいスペクトルは、例えば、それぞれ２０番目の新しいサンプルに対してのみ計算される。２つのスペクトルの間のサンプルにおけるこの距離は、好ましくは、コントローラ６０２によって与えられる。コントローラ６０２は、ＩＦＦＴ処理器（逆ＦＦＴ処理器）６０４に供給するように構成される。ＩＦＦＴ処理器６０４は、重複した操作で作動するように構成される。特に、ＩＦＦＴ処理器６０４は、重複加算操作を実行するために、変更されたスペクトルの大きさと位相に基づいたスペクトルごとに一つのＩＦＦＴを実行することによって、逆短時間フーリエ変換を実行するように構成される。ＩＦＦＴ処理器６０４からは、結果として生じた時間信号が得られる。重複加算操作は、分析窓の効果を排除する。

【００４８】

時間信号の拡張は、２つのスペクトルの間の距離ｂによって達成される。２つのスペクトルは、ＩＦＦＴ処理器６０４によって処理される。２つのスペクトルの間の距離ｂは、ＦＦＴスペクトルの発生におけるスペクトル間の距離ａより大きい。基本的な考え方は、分析ＦＦＴより遠くに離れているＩＦＦＴによって音声信号を拡張することである。その結果、合成音声信号における一時的な変化が、元の音声信号より緩やかに起こる。

【００４９】

しかしながら、ブロック６０６の中で位相が再スケーリングされないと、これは人工物（分散）を導く。例えば、１つの周波数ビンが、４５°で連続した位相値が実行されるために考慮されるとき、これは、このフィルタバンクの中の信号が、１サイクルの１／８の割合で、すなわち、時間間隔あたり４５°で、位相において増加することを含意する。この時間間隔は、連続したＦＦＴの間の時間間隔である。仮に、ＩＦＦＴが、相互により遠くに離れているならば、これは、４５°の位相増加が、より長い時間間隔に渡って起こることを意味する。これは、位相シフトのために、その後の重複加算処理における不一致が起こり、不必要な信号相殺がもたらされることを意味する。この不一致を排除するために、位相は、音声信号が時間内に拡張されたのと同じ係数によって再スケーリングされる

。それぞれのＦＦＴスペクトル値の位相は、係数ｂ／ａによって増加し、その結果、この不一致は排除される。

【００５０】

一方、図５ｃで示された実施形態において、振幅／周波数制御信号の挿入による拡張は、図５ａのフィルタバンクの構成の中の、信号発振器ごとに達成される。図６における拡張は、２つのＦＦＴスペクトルの間の距離ａより長い２つのＩＦＦＴスペクトルの間の距離ｂによって達成される。しかしながら、人工物（分散）防止のために、位相の再スケーリングが、ｂ／ａに従って実行される。

【００５１】

位相音声分析合成装置の詳細な記述に関して、以下の文献が参照される。

（１）「位相音声分析合成装置：チュートリアル」マークダルソン著、コンピュータ音楽ジャーナル、１０巻、Ｎｏ．４、ページ１４～２７、１９８６年（“Ｔｈｅ ｐｈａｓｅ Ｖｏｃｏｄｅｒ：Ａ ｔｕｔｏｒｉａｌ”，Ｍａｒｋ Ｄｏｌｓｏｎ，Ｃｏｍｐｕｔｅｒ Ｍｕｓｉｃ Ｊｏｕｒｎａｌ，ｖｏｌ．１０，ｎｏ．４，ｐｐ．１４－２７，１９８６）

（２）「ピッチシフトのための新しい位相音声分析合成装置の技術、調和および他のエキゾチックな効果」、Ｌ．ラロッシュォおよびＭ．ダルソン著、音声と音響のための信号処理の応用に関する１９９９年ＩＥＥＥ研究集会の会報、ニューパルツ、ニューヨーク、１９９９年１０月１７日～２０日、ページ９１～９４（“Ｎｅｗ ｐｈａｓｅ Ｖｏｃｏｄｅｒ ｔｅｃｈｎｉｑｕｅｓ ｆｏｒ ｐｉｔｃｈ－ｓｈｉｆｔｉｎｇ，ｈａｒｍｏｎｉｚｉｎｇ ａｎｄ ｏｔｈｅｒ ｅｘｏｔｉｃ ｅｆｆｅｃｔｓ”，Ｌ．Ｌａｒｏｃｈｅ ｕｎｄ Ｍ．Ｄｏｌｓｏｎ，Ｐｒｏｃｅｅｄｉｎｇｓ １９９９ ＩＥＥＥ Ｗｏｒｋｓｈｏｐ ｏｎ ａｐｐｌｉｃａｔｉｏｎｓ ｏｆ ｓｉｇｎａｌ ｐｒｏｃｅｓｓｉｎｇ ｔｏ ａｕｄｉｏ ａｎｄ ａｃｏｕｓｔｉｃｓ，Ｎｅｗ Ｐａｌｔｚ，Ｎｅｗ Ｙｏｒｋ，Ｏｃｔｏｂｅｒ １７ － ２０，１９９９，ｐａｇｅｓ ９１ ｔｏ ９４；”）

（３）「瞬間的事象を処理する中間位相音声分析合成装置の新しいアプローチ」、Ａ．ローベル著、デジタル音声効果（ＤＡＦｘ－０３）に関する第６回国際会議の議事録、ロンドン、イギリス（２００３年９月８日～１１日）、ページＤＡＦｘ－１～ＤＡＦｘ－６（Ｎｅｗ ａｐｐｒｏａｃｈｅｄ ｔｏ ｔｒａｎｓｉｅｎｔ ｐｒｏｃｅｓｓｉｎｇ ｉｎｔｅｒｐｈａｓｅ ｖｏｃｏｄｅｒ”，Ａ．Ｒｏｂｅｌ，Ｐｒｏｃｅｅｄｉｎｇ ｏｆ ｔｈｅ ６ｔｈ ｉｎｔｅｒｎａｔｉｏｎａｌ ｃｏｎｆｅｒｅｎｃｅ ｏｎ ｄｉｇｉｔａｌ ａｕｄｉｏ ｅｆｆｅｃｔｓ（ＤＡＦｘ－０３），Ｌｏｎｄｏｎ，ＵＫ，Ｓｅｐｔｅｍｂｅｒ ８－１１，２００３，ｐａｇｅｓ ＤＡＦｘ－１ ｔｏ ＤＡＦｘ－６）

（４）「位相固定された音声分析合成装置」、メラープケット著、音声と音響のための信号処理の応用に関する１９９９年ＩＥＥＥ ＡＳＳＰ研究集会の会報、（“Ｐｈａｓｅ－ｌｏｃｋｅｄ Ｖｏｃｏｄｅｒ”，Ｍｅｌｌｅｒ Ｐｕｃｋｅｔｔｅ，Ｐｒｏｃｅｅｄｉｎｇｓ １９９５，ＩＥＥＥ ＡＳＳＰ，Ｃｏｎｆｅｒｅｎｃｅ ｏｎ ａｐｐｌｉｃａｔｉｏｎｓ ｏｆ ｓｉｇｎａｌ ｐｒｏｃｅｓｓｉｎｇ ｔｏ ａｕｄｉｏ ａｎｄ ａｃｏｕｓｔｉｃｓ，）

（５）米国特許出願Ｎｏ．６，５４９，８８４。

【００５２】

また、信号拡張のための他の方法は、例えば、「ピッチ同期重複加算」法などが利用可能である。ピッチ同期重複加算法（要するにＰＳＯＬＡ法）は、スピーチ信号の記録がデータベースの中に位置している合成方法である。スピーチ信号が周期信号である限り、スピーチ信号は基本周波数（ピッチ）の情報と共に提供される。そして、それぞれの期間の初めが印付けされる。合成において、これらの期間は、窓関数によって、所定の周囲と共に切り取られ、適した位置で音声信号に加算され合成される。望ましい基本周波数がデータベース入り口の周波数より高いか、または、低いかに依存して、スピーチ信号は、元の

スピーチ信号より密度が高いか否かに従って結合される。可聴持続時間を調整するために、期間が２倍に省略されるか、または出力される。この方法はＴＤ－ＰＳＯＬＡ法と称される。ここで、「ＴＤ」は時間領域を表し、ＴＤ－ＰＳＯＬＡ法が時間領域で作動することを強調する。さらなる発展は、多重帯域再合成重複加算（ＭｕｌｔｉＢａｎｄ　Ｒｅｓｙｎｔｈｅｓｉｓ　ＯｖｅｒＬａｐ　Ａｄｄ）法、略してＭＢＲＯＬＡ法である。ここで、データベースの中の構成要素は、前処理で一定の基本周波数とされ、階調音の位相位置は規格化される。これによって、一つの構成要素から次の構成要素への転移の合成において、少ない知覚干渉がもたらされ、達成されるスピーチの品質はより高い。

【００５３】

別の代替において、音声信号は拡張される前に、帯域通過フィルタにかけられる。その結果、拡張されて減衰された後の信号は、既に望ましい部分を含み、その後の帯域通過フィルタリングは省略される。この場合、帯域通過フィルタは、帯域幅拡張の後にフィルタから出力された音声信号の部分が、帯域通過フィルタの出力信号にまだ含まれるように、設定される。その結果、帯域通過フィルタは、拡張されて減衰された後の音声信号に含まれていない周波数領域を含む。この周波数領域をもつ信号は、合成高周波信号を形成する所望の信号である。

【００５４】

図１で示される信号操作器は、さらに、信号線１２１上の非処理の「そのまま」の状態の、または、合成された状態の瞬間的事象をもつ音声信号の別の処理のための信号調整器１３０を含む。この信号調整器１３０は、帯域幅拡張アプリケーションの中の信号減衰器である。信号調整器１３０は、出力にて高帯域信号を発生する。信号調整器１３０は、さらに、ＨＦＲ（高周波再構成）データストリームと共に伝送されべき高周波（ＨＦ）パラメータを使用することによって、元の高帯域信号の特性に密接に類似するように改造することができる。

【００５５】

図７ａと図７ｂは帯域幅拡張シナリオを示す。それは、図７ｂの帯域幅拡張符号器７２０の中の信号調整器１３０の出力信号を有効に使用できる。音声信号は、入力７００にて低帯域通過／高帯域通過の組み合わせフィルタ７０２に送り込まれる。低帯域通過／高帯域通過の組み合わせフィルタ７０２の一方は、低帯域通過（ＬＰ）フィルタを含み、図７ａの符号７０３で示される音声信号７００の低帯域通過フィルタをかけられたバージョンを発生する。この低帯域通過フィルタをかけられた音声信号は、音声符号器７０４で符号化される。音声符号器７０４は、例えば、ＭＰ３符号器（ＭＰＥＧ１　３層）、または、ＡＡＣ符号器、または、ＭＰＥＧ４規格で説明される周知のＭＰ４符号器である。帯域が制限された音声信号７０３の透明な、または、有利に知覚的に透明な表現を提供する二者択一の音声符号器が、符号器７０４の中で使用され、完全に符号化された、または、知覚的に符号化された（好ましくは知覚的に透明に符号化された）音声信号７０５をそれぞれ発生させる。

【００５６】

音声信号の上側の帯域は、組み合わせフィルタ７０２の、「ＨＰ」によって指示された高帯域通過部分の出力７０６にて出力される。音声信号の高帯域部分、すなわち、ＨＦ部分として指示された上側の帯域またはＨＦ帯域は、パラメータ計算機７０７に供給される。パラメータ計算機７０７は、異なるパラメータを計算するように構成されている。これらのパラメータは、例えば、各精神音響周波数グループまたはバーク（Ｂａｒｋ）スケールの各バーク帯域のためのスケール係数の表現による比較的粗い解像度において、出力７０６の上側の帯域のスペクトルエンベロープ（包絡線）である。パラメータ計算機７０７によって計算される別のパラメータは、上側の帯域の雑音床である。帯域あたりのエネルギーは、好ましくは、上側の帯域におけるエンベロープのエネルギーに関係する。パラメータ計算機７０７によって計算される別のパラメータは、上側の帯域の各部分帯域ごとの色調測定を含む。色調測定は、スペクトルエネルギーがこの帯域でどのように分配されるかを示す。すなわち、この帯域におけるスペクトルエネルギーが比較的一様に分配されて

いる（その場合、色調の信号がこの帯域に存在していない）かどうか、または、この帯域におけるスペクトルエネルギーが、帯域の所定の位置に比較的強く集中している（その場合、色調の信号がこの帯域に存在している）かどうか、を示す。

【００５７】

別のパラメータが、その高さとその周波数に関して上側の帯域の中で比較的強く突出するピークを明らかに符号化することの中に存在する。帯域幅拡張概念が、上側の帯域の、際立った正弦波様の部分の明白な符号化をしない再構成において、上側の帯域を非常に粗く回復するだけである。または、上側の帯域を全く回復しない。

【００５８】

どのような場合でも、パラメータ計算機７０７は、上側の帯域のパラメータ７０８だけを発生させるように構成される。パラメータ７０８は、同じエントロピー再生ステップに従属する。エントロピー再生ステップは、量子化されたスペクトル値ごとに、音声符号器７０４の中で、例えば、差分符号化、予測またはハフマン符号化などが実行される。パラメータ７０８および音声信号７０５が、データストリーム形成器７０９に提供される。データストリーム形成器７０９は、出力側データストリーム７１０を提供するように構成される。出力側データストリーム７１０は、一般的に、例えば、ＭＰＥＧ４規格で規格化されたフォーマットに従ったビットストリームである。

【００５９】

本発明に特に適している復号器側が、図７ｂに示される。データストリーム７１０は、データストリーム解読器７１１に入る。データストリーム解読器７１１は、音声信号部分７０５から帯域幅拡張関係パラメータ部分７０８を分離するように構成されている。パラメータ部分７０８は、パラメータ復号器７１２によって復号されて、復号されたパラメータ７１３を得る。これに並行して、音声信号部分７０５は、音声復号器７１４によって復号されて、音声信号を得る。

【００６０】

実施例によって、音声信号１００は、第１の出力７１５を通して出力される。出力７１５にて、小さい帯域幅と、その結果の低品質とをもつ音声信号が得られる。しかしながら、品質改良のために、本発明の帯域幅拡張器７２０は、出力側で、拡張された帯域幅、または、高帯域幅と、その結果の高品質とをもつ音声信号７１２を得るように実行される。

【００６１】

音声信号は、符号器側の状況において制限される帯域に従属し、高品質の音声符号器によって音声信号の下側の帯域だけを符号化することが、ＷＯ９８／５７４３６から知られている。しかしながら、上側の帯域は、上側の帯域のスペクトルエンベロープを再生させる１セットのパラメータで、非常に粗く特徴付けられるだけである。そして、復号器側では、上側の帯域が合成される。このために、階調音の転移が提案される。復号された音声信号の下側の帯域は、フィルタバンクに供給される。下側の帯域のフィルタバンクチャンネルは、上側の帯域のフィルタバンクチャンネルに接続される、または、「修理」される。そして、それぞれの修理された帯域通過信号は、エンベロープ調整に従属させられる。ここで、特殊解析フィルタバンクに属する合成フィルタバンクが、下側の帯域における音声信号の帯域通過信号と、上側の帯域で調和して修理された、下側の帯域のエンベロープ調整された帯域通過信号と、を受信する。合成フィルタバンクの出力信号は、非常に低いデータ信号速度で符号器側から復号器に伝送された、帯域幅に関して拡張された音声信号である。特に、フィルタバンク領域でのフィルタバンク計算と修理は、高い計算努力になる。

【００６２】

ここに提示された方法は、言及した問題を解決する。本発明の方法の目新しさは、既存の方法と対照して、瞬間的事象を含む第１の窓部分が、操作されるべき音声信号から除去されることを含む。さらに、元の音声信号から第２の窓部分（一般に、第１の窓部分と異なる）が付加的に選択され、一時的なエンベロープが瞬間的事象の周囲にできるだけ保存されるように、操作された音声信号に再挿入されることを含む。。この第２の窓部分は、

時間拡張操作によって変更された凹部に正確に収まるように選択される。正確な収まりが、元の瞬間的事象部分の縁で、結果として起こる凹部の縁の最大の相互相関を計算することによって実行される。

【００６３】

したがって、瞬間的事象の主観的な音質は、もはや分散とエコー効果とによって損なわれない。

【００６４】

適した部分を選択するための瞬間的事象の位置の正確な決定は、例えば、適した期間にわたってエネルギーの移動中心計算を使用することで実行される。

【００６５】

時間拡張係数と共に、第１の時間部分のサイズは、第２の時間部分の必要なサイズを決定する。好ましくは、このサイズは、密接に隣接している瞬間的事象の時間間隔が、個々の一時的事象の人間の知覚の閾値以下である場合にだけ、１つ以上の瞬間的事象が、再挿入のために使用される第２の時間部分によって収容されるように、選択されるべきである。

【００６６】

最大の相互相関に従った瞬間的事象の最適な収まりは、時間において、瞬間的事象の元の位置と比べて、わずかなオフセットを必要とする。しかしながら、一時的な前マスキング効果、および、特に後マスキング効果の存在によって、再挿入された瞬間的事象の位置は、正確に元の位置に合致する必要はない。後マスキングの動作の拡張期間のために、正時間方向における瞬間的事象のシフトが好ましい。

【００６７】

標本抽出率が、その後の減衰ステップによって変更されるとき、元の信号部分を挿入することによって、元の信号部分の音色やピッチが変更される。しかしながら、一般に、これは、精神音響の一時的なマスキング機構によって、瞬間的事象自体によって隠される。特に、仮に、整数係数によって拡張が起こるならば、音色がわずかに変わるのみである。瞬間的事象の周囲の外側は、全てのｎ次階調波（ｎ＝拡張係数）のみで占められるからである。

【００６８】

新しい方法を使用して、時間拡張および転移方法によって瞬間的事象を処理している間に結果として生じる人工物（分散、前エコー、後エコー）が、効果的に防止される。重ねられた（可能な色調）信号部分の品質の潜在的損傷が避けられる。

【００６９】

この方法は、音声信号やそれらのピッチの再生速度が変更される、どんな音声アプリケーションに対しても適している。

【００７０】

次に、図８ａ～図８ｅの文脈の中で、好ましい実施形態について議論する。図８ａは、簡単な時間領域の音声サンプル系列と対照して音声信号の表現を示す。図８ａはエネルギーエンベロープ表示を示す。エネルギーエンベロープ表示は、例えば、時間領域サンプル図のそれぞれの音声サンプルが二乗されるとき、得ることができる。特に、図８ａは瞬間的事象８０１を有する音声信号８００を示す。瞬間的事象８０１は、時間が経過するにつれて、エネルギーの急峻な増加と減少とによって特徴付けられる。当然のことながら、瞬間的事象は、エネルギーが所定の高レベルで維持されているときのエネルギーの急峻な増加や、エネルギーが減少の前の所定の時間の間、高レベルにあるときの急峻な減少も含む。瞬間的事象の特異的パターンは、例えば、手拍子や、打楽器によって発生する他の音調である。さらに、瞬間的事象は、大きな音調で演奏を始める楽器の急激な開始である。前記楽器は、一つの所定の帯域または複数の帯域の中に音声エネルギーを、所定の閾時間より下で、かつ、所定の閾レベルより上で提供する。当然のことながら、図８ａにおいて、音声信号８００のエネルギー変動８０２などの他のエネルギー変動は、瞬間的事象として検出されない。瞬間的事象検出器は周知のものであり、文献で広く説明されており、多く

の異なるアルゴリズムが適用される。アルゴリズムは、周波数選択処理、周波数選択処理の結果と閾値との比較、その後の瞬間的事象が存在するか否かの決定、を含む。

【００７１】

図８ｂは窓を付けられた瞬間的事象８０１を示す。実線によって区切られた領域は、描写された窓形状によって重み付けされた音声信号８００から除去される。破線によって示される領域は、処理の後に再び付加される。特に、所定の瞬間的事象時間８０３で発生する瞬間的事象８０１は、音声信号８００から切り取らなければならない。安全策を取って、瞬間的事象８０１だけではなく、いくつかの隣接／近傍サンプルも、元の音声信号８００から切り取られるべきである。したがって、開始時間８０５から停止時間８０６まで広がる第１の時間部分８０４が決定される。一般に、第１の時間部分８０４は、瞬間的事象時間８０３が第１の時間部分８０４の中に含まれるように選択される。図８ｃは、拡張される前の、瞬間的事象８０１の無い音声信号８００を示す。緩やかに減衰する縁部８０７と８０８から認められるように、第１の時間部分８０４が、矩形の適合枠／窓枠によって切り取られるだけでなく、窓化は、音声信号８００の緩やかに減衰する縁部または側部を有することを実行する。

【００７２】

重要なことに、図８ｃは、図１の信号線１０２の音声信号、すなわち、瞬間的事象が除去された音声信号を示す。緩やかに減衰／増加する側部８０７，８０８は、図４の相互フェーダ１２８によって使用されるべきフェードイン領域またはフェードアウト領域を備える。図８ｄは、図８ｃの音声信号８００の拡張された状態、すなわち、音声信号８００が信号処理器１１０によって処理された状態を示す。したがって、図８ｄの信号は、図１の信号線１１１の信号である。拡張操作によって、第１の時間部分８０４は非常に長くなる。したがって、図８ｄの第１の時間部分８０４が第２の時間部分８０９に拡張される。第２の時間部分８０９は、第２の時間部分の開始時間８１０と第２の時間部分の停止時間８１１とを有する。音声信号８００を拡張することによって、側部８０７，８０８も同様に拡張され、その結果、側部８０７′，８０８′の時間長さも同様に拡張される。第２の時間部分８０９の長さの計算が、図４の計算機１２２によって実行されるとき、この拡張は考慮されなければならない。

【００７３】

第２の時間部分８０９の長さが決定されるとすぐに、第２の時間部分８０９の長さに対応する部分が、図８ｂで破線によって示されるように、図８ａで示された元の音声信号から切り取られる。この後、第２の時間部分８０９が図８ｅに入れられる。議論したように、第２の時間部分８０９の開始時間８１２（すなわち、元の音声信号８００の第２の時間部分８０９の第１の境界）と、第２の時間部分８０９の停止時間８１３（すなわち、元の音声信号８００の第２の時間部分８０９の第２の境界）とは、瞬間的事象時間８０３，８０３′に関して必ずしも対称である必要はない。その結果、瞬間的事象８０１は、元の音声信号８００に位置していた瞬間的事象８０１のまさに同じ時間に位置している。代わりに、図８ｂの時間８１２、８１３は、わずかに変更することができる。従って、元の音声信号８００のこれらの境界の信号形状の間に結果として生じる相互相関は、拡張された音声信号８００の対応部分と、できるだけ同様である。その結果、瞬間的事象８０１の実際の時間８０３は、第２の時間部分８０９の中心から、ある程度まで移動できる。第２の時間部分８０９の中心は、第２の時間部分８０９に関して所定の時間を示す符号８０３′によって図８ｅの中に示される。瞬間的事象８０１の実際の時間８０３′は、図８ｂの第２の時間部分８０９に関して、対応する時間８０３から外れる。図４に関係して符号１２６で議論したように、時間８０３に関する時間８０３′への瞬間的事象８０１の正シフトは、前マスキング効果より顕著である後マスキング効果のために好ましい。図８ｅはさらに、重複／転移領域８１３ａ，８１３ｂを示す。相互フェーダ１２８は、瞬間的事象を有さない拡張された音声信号と、瞬間的事象を含む元の音声信号の複製と、の間の相互フェーダを提供する。

【００７４】

図４で示されるように、第２の時間部分８０９の長さを計算するための計算機１２２は、第１の時間部分８０４の長さと拡張係数を受信するように構成される。また、計算機１２２は、全く同じ第１の時間部分の中に含まれるべき隣接瞬間的事象の許容性に関する情報を受信することができる。したがって、この許容性に基づいて、計算機１２２自体は、第１の時間部分８０４の長さを決定する。その後、拡張係数／短縮係数に依存して、第２の時間部分８０９の長さを計算する。

【００７５】

前述のように、信号挿入機１２０の機能は、信号挿入機１２０が図８ｅの間隙に適した領域を置き換えることである。間隙は、元の音声信号から拡張された音声信号において、拡大され、この適した領域に合致する。すなわち、第２の時間部分８０９を、時間８１２と８１３を決定するための相互相関計算を使用して、処理音声信号に合致させる。好ましくは、相互フェード領域８１３ａと８１３ｂにおいて、相互フェード操作も同様に実行する。

【００７６】

図９は、音声信号のためのサイド情報を発生させるための発生装置を示す。発生装置は、瞬間的事象検出が符号器側で実行されて、この瞬間的事象検出に関するサイド情報が計算されて、復号器側を代表する音声信号マニピュレータに伝送されるとき、本発明の文脈の中で使用できる。このために、図２の瞬間的事象検出器１０３と同様の瞬間的事象検出器が、瞬間的事象を含む音声信号を分析するために用いられる。瞬間的事象検出器は、瞬間的事象時間、すなわち、図８ｂの時間８０３を計算して、この瞬間的事象時間をメタデータ計算機１０４′に伝送する。メタデータ計算機１０４′は、図２のフェードアウト／フェードイン計算機１０４と同様の構成である。一般に、メタデータ計算機１０４′は、音声信号出力インタフェース９００に伝送すべきメタデータを計算できる。このメタデータは、瞬間的事象が除去される境界、すなわち、第１の時間部分８０４の境界である、図８ｂの８０５，８０６や、図８ｂの時間８１２，８１３で示された瞬間的事象挿入（第２の時間部分８０９）のための境界や、瞬間的事象時間８０３，８０３′を含む。後者の場合でさえ、音声信号マニピュレータは、すべての必要なデータ、すなわち、瞬間的事象時間８０３に基づいた第１の時間部分データ、第２の時間部分データなどを決定する立場にある。

【００７７】

メタデータ計算機１０４′によって生成したメタデータは、信号出力インタフェース９００に伝送される。信号出力インタフェース９００は、信号、すなわち、伝送または格納のための出力信号を生成する。出力信号は、メタデータだけ、または、メタデータおよび音声信号を含む。後者の場合、メタデータは、音声信号のサイド情報を表す。このために、音声信号は、信号線９０１を通して信号出力インタフェース９００に伝送される。信号出力インタフェース９００によって生成された出力信号は、どんな種類の記憶媒体にも格納でき、音声信号マニピュレータや、瞬間的事象情報を必要とする、いかなる他の装置にも、どんな種類の伝送チャンネルを通しても送信できる。

【００７８】

本発明は、ブロックが実際の、または、論理的なハードウェアの部品を表すブロック図の文脈の中で説明されているけれども、本発明は、コンピュータが実行する方法によっても、実施することができることに注目するべきである。後者の場合、ブロックは対応する方法のステップを表し、これらのステップは、対応する論理的または物理的なハードウェアブロックによって実行される機能を表す。

【００７９】

記載されている実施例は、本発明の原理のために、単に図示するだけである。配置および本願明細書において記載されている詳細の修正および変更は、他の当業者にとって明らかであるものと理解される。従って、現実の特許請求の範囲だけによって制限され、本願明細書において実施例の説明および説明として示される具体的な詳細だけによって制限されないことが意図される。

【００８０】
本発明の方法の所定の実現要求によって、本発明の方法は、ハードウェアまたはソフトウェアの中で実現することができる。実現は、デジタル格納媒体を使用することで実行できる。特に、ディスク、ＤＶＤ、ＣＤは、その上に保存された電子的に読み込み可能な制御信号を有している。それらは、本発明の方法が実行されるように、プログラム可能なコンピュータシステムと協働する。一般に、本発明は、コンピュータプログラム製品がコンピュータ上で動くとき、機械読み込み可能な媒体上に格納されたプログラムコードを有するコンピュータプログラム製品として実行することができる。プログラムコードは、本発明の方法を実行するために操作される。言い換えれば、本発明の方法は、コンピュータプログラムがコンピュータ上で動くとき、少なくとも本発明の方法の１つを実行するためのプログラムコードを有するコンピュータプログラムである。本発明のメタデータ信号は、デジタル格納媒体などのどんな機械読み込み可能な記憶媒体にも格納できる。

【符号の説明】

【００８１】
１００　瞬間的事象信号除去器
１０３　瞬間的事象検出器
１０４’　メタデータ計算機
１０６　サイド情報抽出器
１１０　信号処理器
１１２　周波数選択分析器
１１３　周波数選択処理装置
１２０　信号挿入器
１２８　相互フェード
１３０　信号調整器
８０１　瞬間的事象
８０３　時間位置
８０４　第１の時間部分
８０９　第２の時間部分
８１３ａ　始めの部分
８１３ｂ　終わりの部分
９００　信号出力インタフェース

【図１】

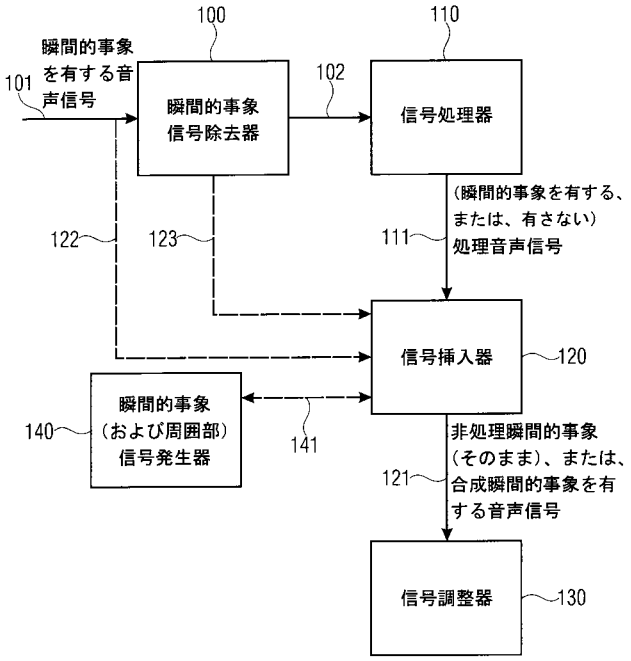



FIGURE 1

【図２】

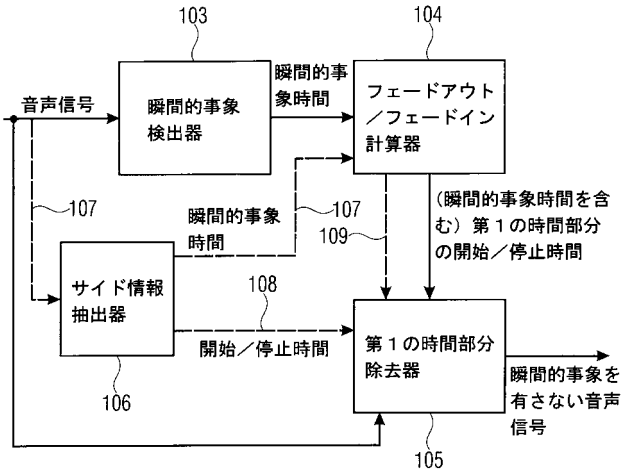



FIGURE 2

【図３Ａ】

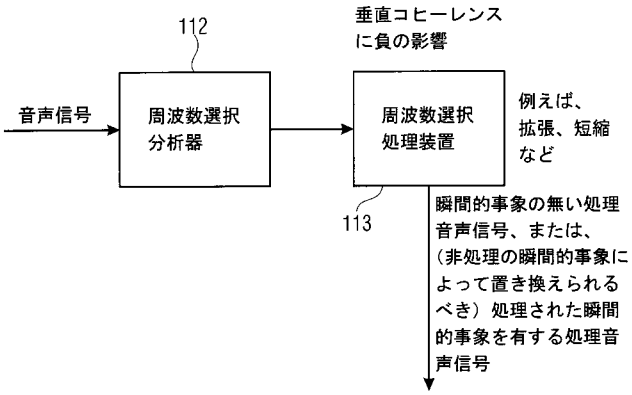



FIGURE 3A

【図３Ｂ】

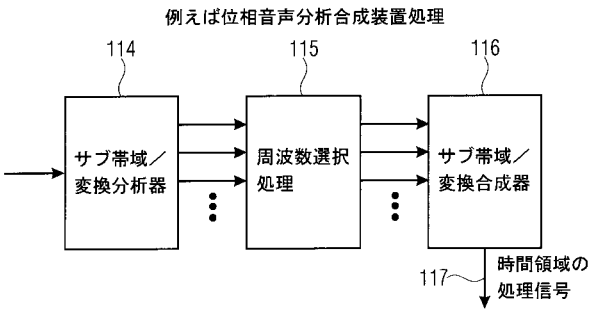



FIGURE 3B

【図４】

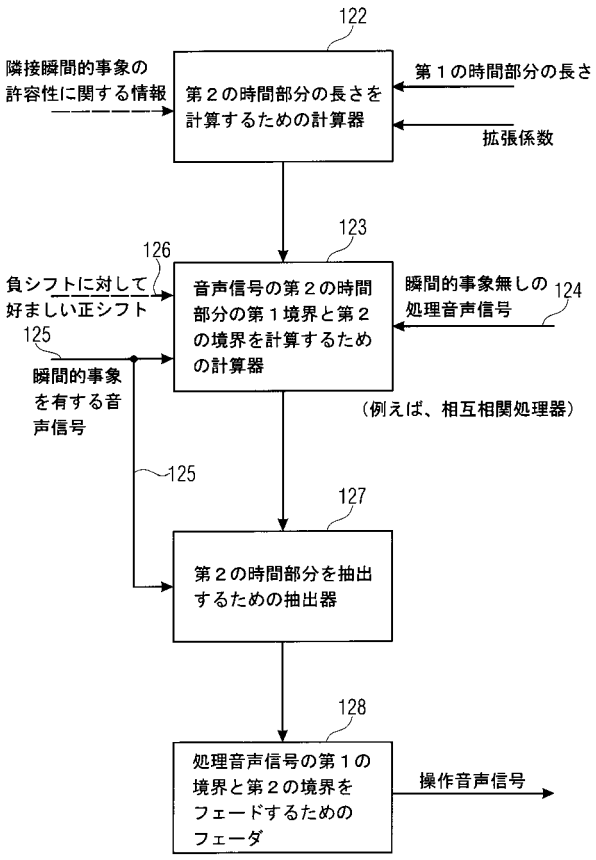
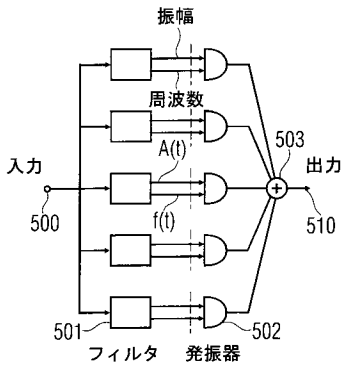



FIGURE 4

【図５Ａ】




スペクトル情報と
時間情報の分離

FIGURE 5A
（フィルタバンクの実施例）

【図５Ｂ】

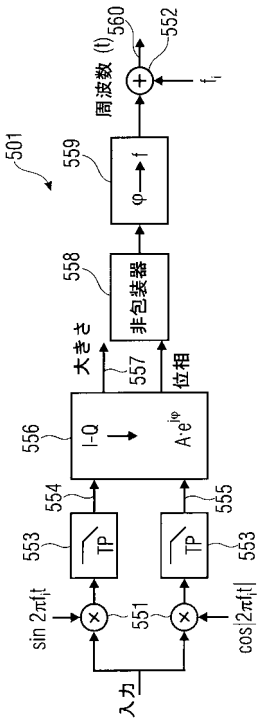
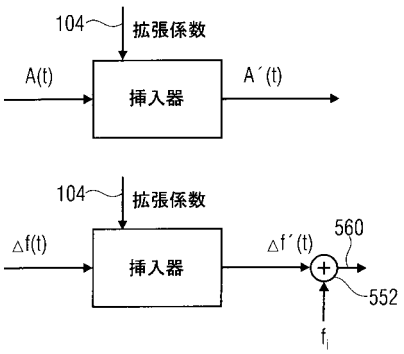



FIGURE 5B

【図５Ｃ】




時間情報の挿入
（拡張）

FIGURE 5C

【図６】

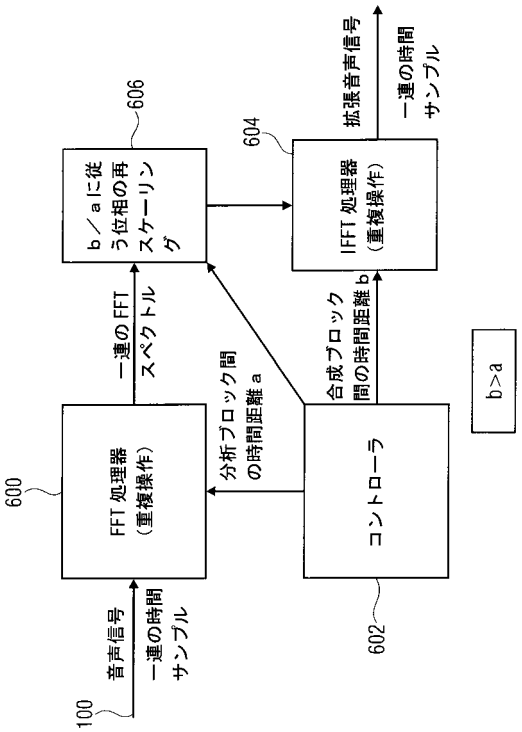



FIGURE 6
（変換実施例）

【図７Ａ】

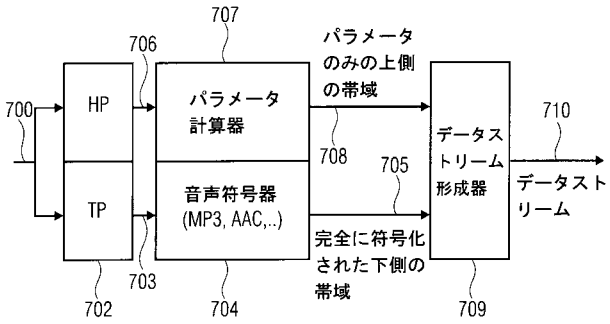



FIGURE 7A
（符号器側）

【図７Ｂ】

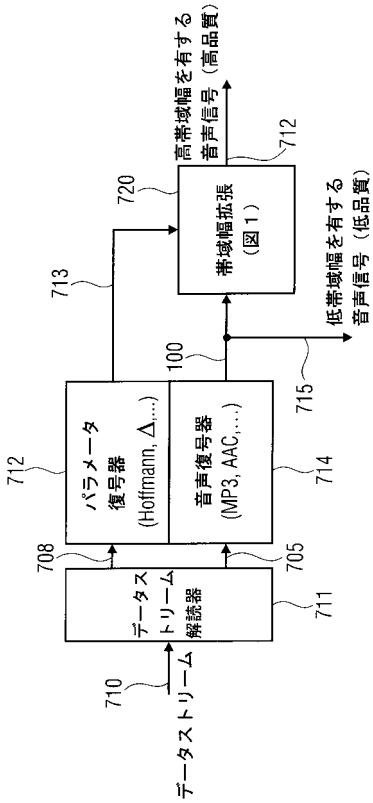



FIGURE 7B
（復号器側）

【図８Ａ】

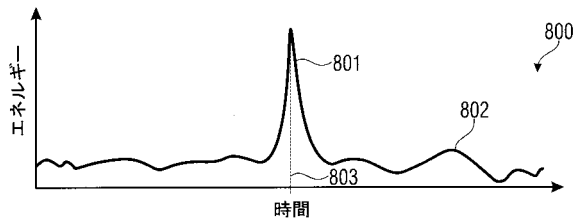



FIGURE 8A

【図８Ｂ】

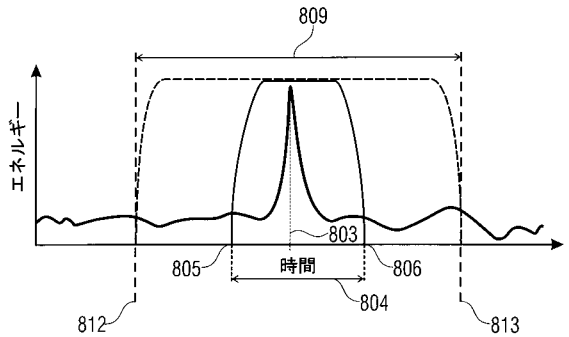



FIGURE 8B

【図８Ｃ】

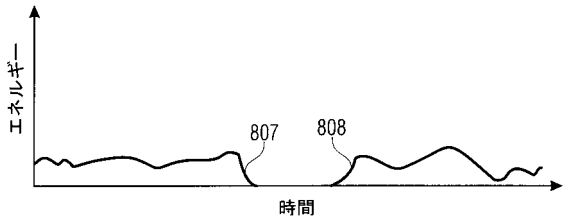



FIGURE 8C

【図８Ｄ】

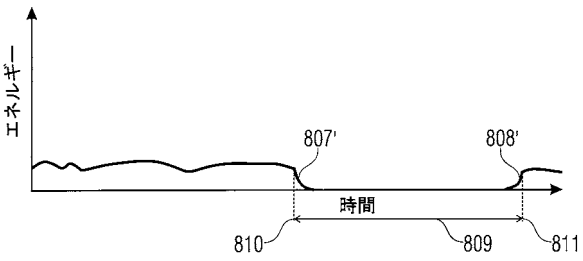



FIGURE 8D

【図８Ｅ】

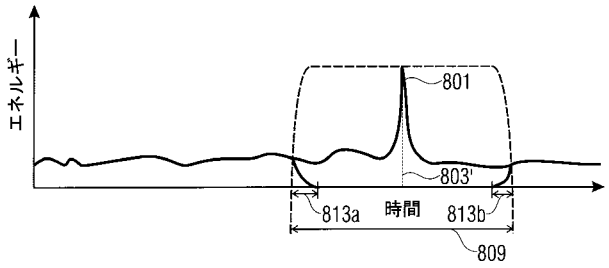



FIGURE 8E

【図９】

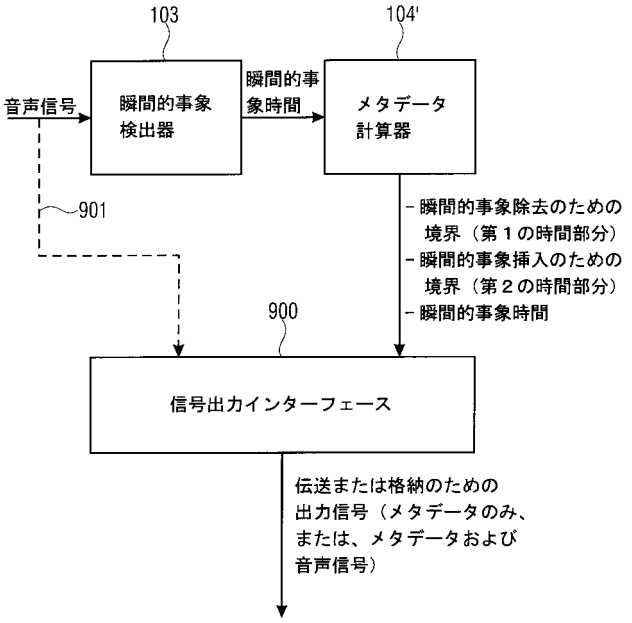



FIGURE 9

フロントページの続き

(72)発明者　レッテルバッハ　ニコラウス
ドイツ連邦共和国　９０４２７　ニュルンベルグ　アーモルバッハー　シュトラーセ　２アー
(72)発明者　ムルトルス　マルクス
ドイツ連邦共和国　９０４６９　ニュルンベルグ　エッツラウブヴェーク　７
(72)発明者　フックス　ギヨーム
ドイツ連邦共和国　９０４０９　ニュルンベルグ　パークシュトラーセ　１２

【外国語明細書】
2012141631000001.pdf
2012141631000002.pdf
2012141631000003.pdf
2012141631000004.pdf