(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5013822号
(P5013822)
(45)発行日　平成24年8月29日(2012.8.29)　(24)登録日　平成24年6月15日(2012.6.15)

(51)Int.Cl.　F I
G10L 19/00 (2006.01)　G10L 19/00 330B
G10L 19/00 330C

請求項の数 11　(全 11 頁)

(21)出願番号　特願2006-304408(P2006-304408)
(22)出願日　平成18年11月9日(2006.11.9)
(65)公開番号　特開2008-122533(P2008-122533A)
(43)公開日　平成20年5月29日(2008.5.29)
審査請求日　平成21年11月6日(2009.11.6)

(73)特許権者　000001007
キヤノン株式会社
東京都大田区下丸子３丁目３０番２号
(74)代理人　100076428
弁理士　大塚　康徳
(74)代理人　100112508
弁理士　高柳　司郎
(74)代理人　100115071
弁理士　大塚　康弘
(74)代理人　100116894
弁理士　木村　秀二
(72)発明者　黒木　大輔
東京都大田区下丸子３丁目３０番２号　キヤノン株式会社内

最終頁に続く

(54)【発明の名称】音声処理装置とその制御方法、及び、コンピュータプログラム

(57)【特許請求の範囲】
【請求項１】
時間方向に分割された音声データを含む音声パケットに基づき、該音声パケットに含まれる音声データの音声処理を行う音声処理装置であって、
第１の期間の音声データを含み、前記第１の期間の直前の期間である第２の期間の音声データ、及び、前記第１の期間の直後の期間である第３の期間の音声データを含まない第１の音声パケットを入力する入力手段と、
前記入力手段により入力された第１の音声パケットに、前記第２の期間の音声データ及び前記第３の期間の音声データの少なくともいずれかの音声データを、該第１の期間の音声データを前記音声処理する際の参照用の音声データとして付加して第２の音声パケットを生成する第１のパケット生成手段と、
前記第２の音声パケットに含まれる、前記第１の期間の音声データ及び前記参照用の音声データを用いて、周波数特性の解析を行い、前記周波数特性の解析結果に基づき前記第１の期間の音声データの前記音声処理を行う音声処理手段と
を備えることを特徴とする音声処理装置。
【請求項２】
前記音声処理手段による前記音声処理が行われた場合に、前記第２の音声パケットから前記参照用の音声データを削除して、該音声処理後の第１の期間の音声データを含む第３の音声パケットを生成する第２のパケット生成手段と、
前記第３の音声パケットに含まれる前記音声処理後の第１の期間の音声データを用いて

音声出力を行う音声出力手段とをさらに備えることを特徴とする請求項１に記載の音声処理装置。
【請求項３】
前記第１のパケット生成手段は、前記第１の期間の音声データの直前及び直後の少なくともいずれかの音声データの期間を、前記分割された音声データの分割単位に基づいて決定することを特徴とする請求項１または２に記載の音声処理装置。
【請求項４】
前記第１のパケット生成手段は、前記参照用の音声データの有無を表す情報を含むように、前記第２の音声パケットを生成することを特徴とする請求項１乃至３のいずれか１項に記載の音声処理装置。
【請求項５】
前記第１のパケット生成手段は、前記参照用の音声データのデータ長を表す情報を含むように、前記第２の音声パケットを生成することを特徴とする請求項１乃至４のいずれか１項に記載の音声処理装置。
【請求項６】
時間方向に分割された音声データを含む第１の音声パケットに基づき、該音声パケットに含まれる音声データの音声処理を行う音声処理装置の制御方法であって、
入力手段が、第１の期間の音声データを含み、前記第１の期間の直前の期間である第２の期間の音声データ、及び、前記第１の期間の直後の期間である第３の期間の音声データを含まない第１の音声パケットを入力する入力工程と、
第１のパケット生成手段が、前記入力工程において入力された第１の音声パケットに、前記第２の期間の音声データ、及び、前記第３の期間の音声データの少なくともいずれかの音声データを、該第１の期間の音声データを前記音声処理する際の参照用の音声データとして付加して第２の音声パケットを生成する第１のパケット生成工程と、
音声処理手段が、前記第２の音声パケットに含まれる、前記第１の期間の音声データ及び前記参照用の音声データを用いて、周波数特性の解析を行い、前記周波数特性の解析結果に基づき前記第１の期間の音声データの前記音声処理を行う音声処理工程と
を備えることを特徴とする音声処理装置の制御方法。
【請求項７】
第２のパケット生成手段が、前記音声処理工程における前記音声処理が行われた場合に、前記第２の音声パケットから前記参照用の音声データを削除して、該音声処理後の第１の期間の音声データを含む第３の音声パケットを生成する第２のパケット生成工程と、
音声出力手段が、前記第３の音声パケットに含まれる前記音声処理後の第１の期間の音声データを用いて音声出力を行う音声出力工程と
をさらに備えることを特徴とする請求項６に記載の音声処理装置の制御方法。
【請求項８】
前記第１のパケット生成工程では、前記第１の期間の音声データの直前及び直後の少なくともいずれかの音声データの期間を、前記分割された音声データの分割単位に基づいて決定することを特徴とする請求項６または７に記載の音声処理装置の制御方法。
【請求項９】
前記第１のパケット生成工程では、前記参照用の音声データの有無を表す情報を含むように、前記第２の音声パケットを生成することを特徴とする請求項６乃至８のいずれか１項に記載の音声処理装置の制御方法。
【請求項１０】
前記第１のパケット生成工程では、前記参照用の音声データのデータ長を表す情報を含むように、前記第２の音声パケットを生成することを特徴とする請求項６乃至９のいずれか１項に記載の音声処理装置の制御方法。
【請求項１１】
コンピュータを、請求項１乃至５のいずれか１項に記載の音声処理装置の各手段として機能させるための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、音声処理装置とその制御方法、及び、コンピュータプログラムに関する。
【背景技術】
【０００２】
放送を受信し映像と音声を出力するテレビ、ＭＰ３やＡＣＣなどの音声圧縮方式で圧縮された音声データを再生する携帯プレーヤー、人が発する声の命令を聞き動作するロボットなど、音声を扱うシステムは、様々な分野に数多く存在する。
【０００３】
それら音声を扱うシステムでは、音声データをハードウエアによる音声処理や、ＤＳＰやプロセッサ上で動作するプログラムであるソフトウエアによる音声処理を行っている。音声処理の例としては、音声出力レベルを一定の範囲に抑えるレベル制御や人の声だけを抽出するためのノイズ除去などがあり、ニュース番組などにおいて、アナウンサーの声を聞き取りやすいよう音声処理を行っているテレビ等がある。その他にも、音声合成、話者認識など様々な音声処理が存在する。
【０００４】
それら音声処理を行っているシステム、即ち、音声処理システムでは、一連の音声入力データを、時間区間もしくはデータサイズごとに分割しパケット化した音声パケットで、音声データを扱うことがある。例えば、ＭＰＥＧ１による圧縮ビデオ・オーディオストリームにおける１音声パケットは、１３Ｂｙｔｅのパケットヘッダと２２７９Ｂｙｔｅの音声データから構成されている。即ち、圧縮した音声データを２２７９Ｂｙｔｅごとに分割し、パケット化している。
【０００５】
図４および図５を用いて、従来例を説明する。図４は、従来の音声パケットの構成を示す図である。
【０００６】
図４において、音声パケット４００は、以下のように構成される。まず、パケットスタート４０１はパケットの開始を示す。パケットタイプ４０２は、パケットの種類・属性を示す。データ長４０３は、音声データのデータ長を示す。そして、音声入力データ４０４は、音声処理を行い出力する区間に対応する音声データである。
【０００７】
このうち、パケットタイプ４０２は、８ビット（０ビット目から７ビット目まで）で構成され、２ビット目に音声データか否かを示すフラグ領域４０２ａを有する。ここでは、フラグ値が１の場合に音声データであることを示す。よって、図４の場合、２ビット目が１であるので、音声データとなる。なお、０ビット目及び１ビット目の領域４０２ｂ及び４０２ｃは、今後の拡張のために予約されたリザーブ領域である。
【０００８】
次に、図５は、従来の音声パケット化の例を示す図である。５０１は、音声波形データであり、横軸が時間、縦軸が出力になる。５０２、５０３、５０４は、音声波形データ５０１を、所定の単位時間で分割しパケット化した音声パケットに含まれる音声入力データである。
【０００９】
音声波形データ５０１は、時間の早い順に右から時間ＴからＴ＋１、Ｔ＋２・・・とＴ＋１０まで１０区間に分割され、パケット化の単位時間を２区間ごとする。よって、音声入力データ５０２は、Ｔ＋２からＴ＋４までの区間をパケット化した音声パケットＮ＋１に含まれる音声入力データとなる。また、音声入力データ５０３は、Ｔ＋４からＴ＋６までの区間をパケット化した音声パケットＮ＋２に含まれる音声入力データとなる。更に、音声入力データ５０４は、Ｔ＋６からＴ＋８までの区間をパケット化した音声パケットＮ＋３に含まれる音声入力データとなる。

【００１０】
これらの音声入力データを処理するための音声処理方法としては、上記のように様々な方法が考案されているが、それらの多くは周波数特性の解析に基づいている。音声入力データに時間区間ごとにＦＦＴ（Ｆａｓｔ　Ｆｏｕｒｉｅｒ　Ｔｒａｎｓｆｏｒｍ：高速フーリエ変換）を行い、周波数成分を抽出し、その解析結果をもとに様々な音声処理を適用する。ＦＦＴでは、２のべき乗個の有限データが連続的に繰り返されているとみなして計算が行われる。このため、音声入力データの時間区間の両端でデータが不連続となる場合、解析結果が不正確となる。従って、１音声パケットが持つ音声入力データが定常とみなせる時間区間でない場合は、前後の音声入力データをもつ音声パケットをメモリ等に蓄積・バッファリングし、それら音声入力データを参照してＦＦＴを適用する。

【００１１】
図５を用いてその一例を説明する。ここでは、音声パケットＮ＋２に含まれる音声入力データ５０３に対して音声処理を行う場合を考える。

【００１２】
音声入力データ５０３の両端は不連続で、定常とみなせる時間区間でない。対象となる音声入力データ５０３と時間的に前後となる、音声パケットＮ＋１及びＮ＋３のそれぞれに含まれる音声入力データ５０２及び５０４が参照用の音声入力データとして必要となる。そこで、音声処理システムでは、これらの参照用の音楽入力データをバッファなどに蓄積しておき、３音声パケット分の音声入力データを参照して音声解析を行い、音声入力データ５０３に対して音声処理を適用し出力する。

【００１３】
なお、特許文献１は、通信時における音声パケットのロスに対して、パケットに冗長データを持たせ、データを復元する発明を記載する。具体的に、冗長データとして排他的論理和のデータを付加している。

【特許文献１】特開２００２－２６１８１９号公報

【発明の開示】

【発明が解決しようとする課題】

【００１４】
音声処理では、負荷の大きな音声処理をリアルタイムで行うため、複数の音声処理モジュールを用いて音声処理を分散して処理を行う場合がある。また、複数チャンネルの音声データを１つの音声処理モジュールを用いて音声処理を並列して処理する場合もある。このように当該音声パケット単位でのみ音声処理を行う場合、音声処理対象となる音声入力データの前後の参照用音声入力データを蓄積できず、精度の高い音声解析が行えないという問題がある。

【００１５】
当該音声パケットの音声入力データのみを参照し音声解析しても、不正確な解析結果を招くこととなり、これに基づく音声処理では音声パケット間の音声出力データの境界で、音声を聞く者に違和感を与えることになる。

【００１６】
そこで、本発明は、単一の音声パケットに含まれる音声入力データのみ用いて音声処理を行う場合でも、精度の高い音声処理を可能とすることを目的とする。

【課題を解決するための手段】

【００１７】
上記課題を解決するための本発明は、時間方向に分割された音声データを含む音声パケットに基づき、該音声パケットに含まれる音声データの音声処理を行う音声処理装置であって、
第１の期間の音声データを含み、前記第１の期間の直前の期間である第２の期間の音声データ、及び、前記第１の期間の直後の期間である第３の期間の音声データを含まない第１の音声パケットを入力する入力手段と、
前記入力手段により入力された第１の音声パケットに、前記第２の期間の音声データ及

び前記第３の期間の音声データの少なくともいずれかの音声データを、該第１の期間の音声データを前記音声処理する際の参照用の音声データとして付加して第２の音声パケットを生成する第１のパケット生成手段と、

前記第２の音声パケットに含まれる、前記第１の期間の音声データ及び前記参照用の音声データを用いて、周波数特性の解析を行い、前記周波数特性の解析結果に基づき前記第１の期間の音声データの前記音声処理を行う音声処理手段とを備える。

【発明の効果】

【００１８】

本発明によれば、単一の音声パケットに含まれる音声入力データのみ用いて音声処理を行う場合でも、精度の高い音声処理を行うことができる。

【発明を実施するための最良の形態】

【００１９】

以下、添付図面を用いて、本発明における実施形態を説明する。

【００２０】

図１は、本発明における音声パケットの構成例を示す図である。図１において、音声パケット１００は、以下のように構成される。まず、パケットスタート１０１はパケットの開始を示す。パケットタイプ１０２は、パケットの種類・属性を示す。データ長１０３は、音声データのデータ長を示す。そして、音声入力データ１０４は、音声処理を行い出力する区間に対応する音声データである。ここまでは、図４で説明した従来の音声パケットの構成と同様である。

【００２１】

そして、本実施形態に対応する音声パケットは、この構成に追加して、参照用のデータ領域１０５及び１０６を含む点に特徴を有する。まず、データ領域１０５は、音声処理を行い出力する区間に対応する音声入力データ１０４に対して、時間的に直前の参照用の音声入力データ１０５ｂとそのデータ長１０５ａとを含む。次に、データ領域１０６は、音声処理を行い出力する区間に対応する音声入力データ１０４に対して、時間的に直後の参照用の音声入力データ１０６ｂとそのデータ長１０６ａとが含まれる。

【００２２】

また、本実施形態に対応する音声パケットでは、パケットタイプ１０２に含まれる情報に更なる特徴を有する。まずパケットタイプ１０２は、８ビット（０ビット目から７ビット目）までで構成され、２ビット目のフラグ領域１０２ａに音声データか否かを示すフラグを有する。ここまでは、従来と同様である。その上で本実施形態では、０ビット目及び１ビット目という従来はリザーブ領域とされていた領域に、参照用の音楽入力データがパケット内に付加されているか否か（存在の有無）を示すフラグ領域１０２ｂ及び１０２ｃを設定している。

【００２３】

具体的に、１ビット目のフラグ領域１０２ｂには、データ領域１０５に参照用の音声入力データ１０５ｂが含まれるか否かを表す情報としてのフラグ値が設定される。また、０ビット目のフラグ領域１０２ｃには、データ領域１０６に参照用の音楽入力データ１０６ｂが含まれるか否かを表す情報としてのフラグ値が設定される。これらのフラグ値が「１」の場合に、対応する領域にデータを含むこととなる。本実施形態では、２ビット目から０ビット目までの各フラグ領域のフラグ値が１であるので、領域１０４から１０６までに、音楽入力データをそれぞれ含むこととなる。

【００２４】

なお、本実施形態において、参照用の音声入力データは、音声入力データ１０４に対して時間的に直前の音声入力データ、直後の音声入力データ、もしくは両方の音声入力データのいずれでもよい。

【００２５】

次に、図２を参照して、音声データのパケット化について説明する。音声データ２０１は、図５の音声データ５０１と同一の音声データである。図２において、横軸が時間、縦

軸が出力になる。音声データ２０２及び２０６は、本実施形態に対応する音声パケット１００に含まれる音声データを示す。まず、音声データ２０２において、２０４は、音声パケットＮ＋１における音声処理を行い出力する区間に対応する音声入力データを示す。また、２０３は、音声入力データ２０４に対して時間的に直前の参照用の音声入力データを示す。２０５は、音声入力データ２０４に対して時間的に直後の参照用の音声入力データを示している。

【００２６】

次に、音声データ２０６において、２０８は、音声パケットＮ＋２における音声処理を行い出力する区間に対応する音声入力データを示す。次に、２０７は、音声入力データ２０８に対して時間的に直前の参照用の音声入力データを示す。２０９は、音声入力データ２０８に対して時間的に直後の参照用の音声入力データを示す。

【００２７】

図２において、音声データ２０１は、時間の早い順に右から時間ＴからＴ＋１、Ｔ＋２・・・とＴ＋１０まで、時間方向において１０区間に分割される。本実施形態では、分割された音声データ２０１を２区間ごとの単位で音声入力データとしてパケット化する。また、パケット化された音声入力データの前後の音声データ１区間ずつを、参照用の音声入力データとしてデータ領域１０５及び１０６に付加する。

【００２８】

例えば、Ｔ＋２からＴ＋４までの区間の音声データを音声入力データ１０４としてパケット化して音声パケットＮ＋１（２０２）を構成する場合、該音声パケット２０２には以下の音声入力データが含まれる。即ち、Ｔ＋２からＴ＋４までの音声入力データ２０４、Ｔ＋１からＴ＋２までの直前の参照用の音声入力データ２０３、Ｔ＋４からＴ＋５までの直後の参照用の音声入力データ２０５である。

【００２９】

次に、Ｔ＋４からＴ＋６までの区間の音声データを音声入力データ１０４としてパケット化して音声パケットＮ＋２（２０６）を構成する場合、該音声パケット２０６には以下の音声入力データが含まれる。即ち、Ｔ＋４からＴ＋６までの音声入力データ２０８、Ｔ＋３からＴ＋４までの直前の参照用の音声入力データ２０７、Ｔ＋６からＴ＋７までの直後の参照用の音声入力データ２０９である。

【００３０】

音声パケットＮ＋１（２０２）に対して音声処理を行う場合、まずは、音声入力データ２０４と、参照用の音声入力データ２０３及び２０５とを用いてＦＦＴによる音声解析を行う。そして、この解析結果を用いて、対象となる音声入力データ２０４に対する音声処理を行う。なお、音声処理の例としては、周波数特性の解析を利用した上述のレベル制御、ノイズ除去、音声合成、或いは、話者認識などが含まれるが、ＦＦＴによる音声解析も音声処理の一部に含むことができる。

【００３１】

このようにして、音声入力データ２０４単独ではなく、参照用の音声入力データ２０３及び２０５を併せて用いて音声解析を行うことにより、より精度の高い解析が可能となる。また、音声入力データ２０４に対して、適切なパラメータを用いた音声処理が可能となる。

【００３２】

なお、音声データ２０１の先頭では、時間的に直前の音声入力データが存在しないため、直後の参照用の音声入力データのみを付加する。また、音声データ２０１の最後尾では、時間的に直後の音声入力データが存在しないため、直前の参照用の音声入力データのみを付加する。

【００３３】

また、図２では、参照用の音声入力データとして、直前、直後の１単位時間分の音声データを利用したが、発明の実施形態はこれに限定されるものではない。即ち、分割された音声データの任意の分割単位に基づいて選択することができ、たとえば、分割単位の２つ

分の音声データを参照用の音声入力データとして利用してもよいし、更に多くの分割単位分の音声データを利用してもよい。

【００３４】

次に、図３を用いて、本実施形態に対応する音声処理装置の構成例を説明する。なお、該音声処理装置は、対応する処理プログラムをＣＰＵのような演算装置により実行することによって実現されてもよいし、専用のハードウエアロジック回路やＤＳＰを用いて実現してもよい。

【００３５】

図３において、音声パケット群３０１が音声処理装置に入力される。音声パケット群３０１に含まれる各音声パケットは、図４及び図５に示すようなパケット構造を有し、音声データをある一定時間区間で分割し、分割した音声データとヘッダ情報とをパケット化した音声パケットである。本実施形態では、時間の早い順Ｎ、Ｎ＋１、Ｎ＋２・・・に入力される。なお、この時点では、音声処理の為の参照用の音声入力データは含まれていない。

【００３６】

この音声パケット群３０１は、参照音声入力データ付加モジュール３０２に入力される。しかし、上述のように、音声パケット群３０１に含まれる各音声パケットの有する音声入力データのみでは、精度の高い音声処理が行えない。そこで、参照音声入力データ付加モジュール３０２は、時間的に前後の音声パケットに含まれる音声入力データを参照用の音声入力データとして当該音声パケットに付加し、また、参照用の音声入力データが付加されたことを示すフラグを立てる。これにより、新たな音声パケット３０３ａ乃至３０３ｃを生成する。

【００３７】

参照音声入力データ付加モジュール３０２からは、３つの音声処理モジュール３０４ａ乃至３０４ｃに対して並列に音声パケット３０３ａ乃至３０３ｃの出力が行われる。

【００３８】

例えば、音声パケット群３０１におけるパケット番号Ｎの音声パケットについては、該音声パケットの音声入力データＮに対し時間的に直前の参照用の音声入力データとして、パケット番号Ｎ－１の音声パケットの音声入力データＮ－１が付加される。また、該音声パケットの音声入力データに対し時間的に直後の参照用の音声入力データとして、パケット番号Ｎ＋１の音声パケットの音声入力データＮ＋１が付加される。

【００３９】

なお、図３では、音声パケット３０３ａ乃至３０３ｃについて、３つのパケット番号のセットとして記載しているが、これは各番号に対応する音声入力データを含む１つのパケットを意味するものとして記載している。たとえば、音声パケット３０３ａであれば、パケット番号Ｎ－１、Ｎ、Ｎ＋１の３つの音声パケットに含まれる音声入力データＮ－１、Ｎ、Ｎ＋１を含む１つのパケットを意味する。即ち、各パケットにおいて、３つの番号の中心に位置する番号が、該パケットにおける音声入力データを示し、残りの２つが参照用の音声入力データを示している。

【００４０】

音声処理モジュール３０４ａ乃至３０４ｃは、入力された各音声パケットの音声処理を行う。音声モジュール３０４ａ乃至３０４ｃのそれぞれは、音声パケット３０３ａ乃至３０３ｃを受け取ると、図２のパケット構造に示すフラグ領域１０２ｂと、フラグ領域１０２ｃとにおけるフラグ値を確認する。このフラグ値により、音声パケット３０３ａ等に参照用の音声入力データが含まれているか否かを判定することができる。もし、参照用の音声入力データが付加されている場合は、パケット内に含まれる音声入力データと、参照用の各音声入力データとを用いて音声解析を行い、その結果に基づき音声入力データに対して音声処理を行う。音声処理を行った後、参照用の音声入力データを音声パケット内から削除して、音声処理後の音声入力データを領域１０４に含む音声パケット３０５ａ乃至３０５ｃを出力する。

【００４１】
より具体的に、たとえば、音声処理モジュール３０４ａは、音声パケット３０３ａの音声入力データＮと参照用の音声入力データＮ－１及びＮ＋１とを用いて音声解析を行い、音声入力データＮに対して処理を行い、音声入力データＮ'とする。その後、参照用の音声入力データＮ－１及びＮ＋１を削除して、音声パケット３０５ａを出力する。
【００４２】
音声処理モジュール３０４ｂ及び３０４ｃでも同様にして、音声入力データＮ＋１’及び音声入力データＮ＋２’をそれぞれ含む音声パケット３０５ｂ及び３０５ｃをそれぞれ出力する。
【００４３】
音声出力モジュール３０６は、音声処理モジュールから出力された音声パケット３０５ａ乃至３０５ｃに基づき、音声出力を行う。音声出力の順は、パケット番号（Ｎ’、Ｎ＋１’、Ｎ＋２’）に従う。音声３０７は、音声出力モジュールから出力された音声である。
【００４４】
なお、図３では、参照用の音声入力データとして、パケット化されている直前、直後の２単位時間分の音声データを利用したが、発明の実施形態はこれに限定されるものではない。即ち、分割された音声データの任意の分割単位に基づいて選択することができ、たとえば、分割単位の１つ分の音声データを参照用の音声入力データとして利用してもよいし、更に多くの分割単位分の音声データを利用してもよい。
【００４５】
以上のようにして、参照用の音声入力データを音声パケットに付加することで、分散処理時に他の音声パケットを参照することなく音声処理を行う場合に、精度の高い音声処理を行うことができる。これは、分散処理のみではなく、並列処理など他の音声パケットを参照することなく音声パケット単位で音声処理を行う装置においても、同様に精度の高い音声処理を行うことが可能となる。
【００４６】
また、音声パケットが参照用の音声入力データを有するか否かを表す情報（フラグ値）が与えられるので、音声パケットに参照用の音声入力データが付加されているか否かを、当該音声パケットを解析するだけで知ることができる。
【００４７】
更に、参照用の音声入力データのデータ長の情報に基づいて、該参照用の音声入力データを音声パケットから効率的に抽出することができる。
【００４８】
［その他の実施形態］
本発明の目的は、前述した機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムに供給し、そのシステムがプログラムコードを読み出し実行することによっても達成される。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現し、そのプログラムコードを記憶した記憶媒体は本発明を構成する。また、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム（ＯＳ）などが実際の処理の一部または全部を行い、その処理によって前述した機能が実現される場合も含まれる。
【００４９】
さらに、以下の形態で実現しても構わない。すなわち、記憶媒体から読み出されたプログラムコードを、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込む。そして、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるＣＰＵなどが実際の処理の一部または全部を行って、前述した機能が実現される場合も含まれる。
【００５０】
本発明を上記記憶媒体に適用する場合、その記憶媒体には、先に説明したフローチャー

トに対応するプログラムコードが格納されることになる。

【図面の簡単な説明】

【００５１】

【図１】発明の実施形態に対応する音声パケットの構成例を示す図である。

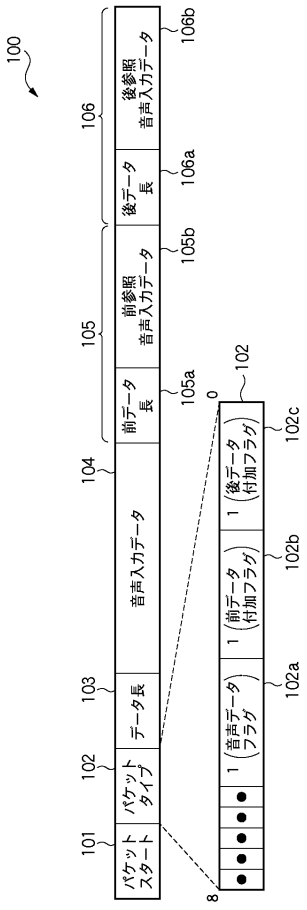
【図２】発明の実施形態に対応する音声パケットの生成例を説明するための図である。

【図３】発明の実施形態に対応する分散処理による音声処理装置の構成例を示す図である。

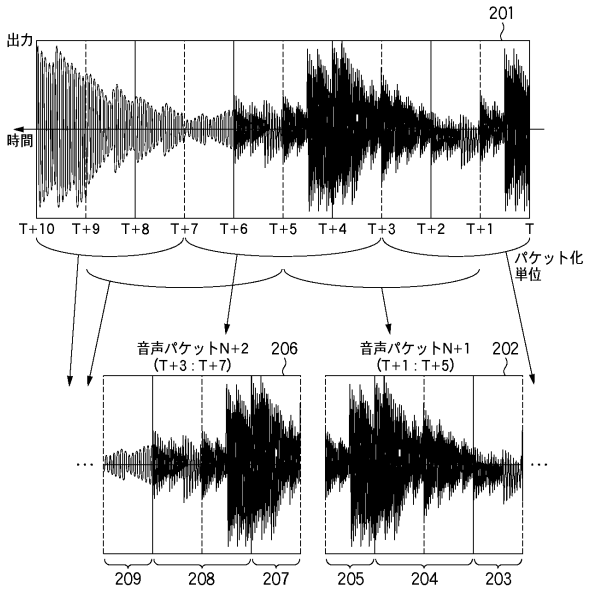
【図４】従来例における音声パケットの構成を示す図である。

【図５】従来例における音声パケット化の例を示す図である。

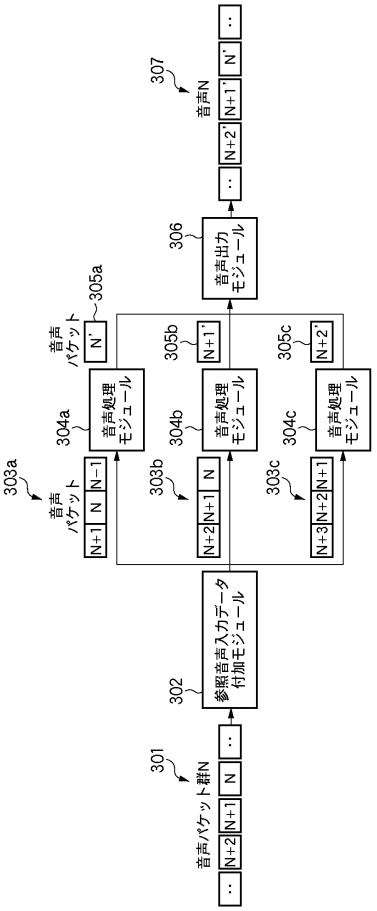
【図１】



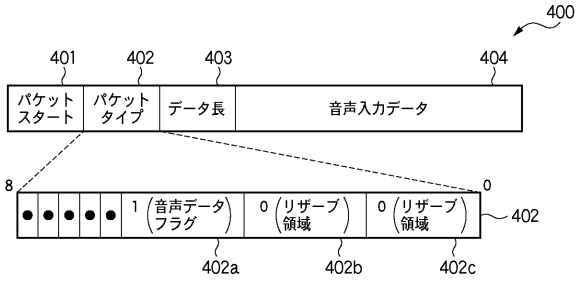
【図２】



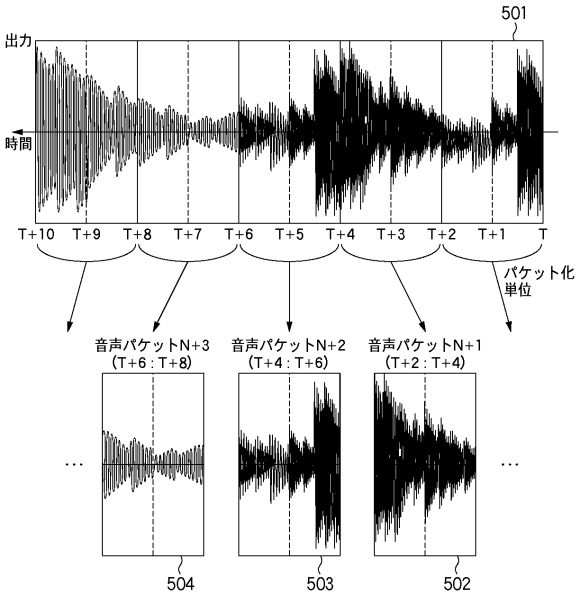
【図３】


307
音声N
306
音声出力モジュール
305a
305b
305c
音声パケット
304a
304b
304c
音声処理モジュール
303a
303b
303c
302
参照音声入力データ付加モジュール
301
音声パケット群N

【図４】


400
401
402
403
404
パケットスタート
パケットタイプ
データ長
音声入力データ
8
0
1 (音声データフラグ)
0 (リザーブ領域)
0 (リザーブ領域)
402a
402b
402c

【図５】


501
出力
時間
T+10
T+9
T+8
T+7
T+6
T+5
T+4
T+3
T+2
T+1
T
パケット化単位
音声パケットN+3 (T+6 : T+8)
音声パケットN+2 (T+4 : T+6)
音声パケットN+1 (T+2 : T+4)
504
503
502

フロントページの続き

審査官　毛利　太郎

(56)参考文献　特開２００６－１４６２４７（ＪＰ，Ａ）
特開２００６－１５３９０８（ＪＰ，Ａ）
特開２００４－３５６８９８（ＪＰ，Ａ）
特開２００２－２６１８１９（ＪＰ，Ａ）
特開２００１－１８４０８０（ＪＰ，Ａ）
国際公開第２００６／０９０５８９（WO，Ａ１）
特表２００３－５３３９１６（ＪＰ，Ａ）
特開２００８－１１１９９１（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１９／００－１９／１４
Ｈ０３Ｍ　７／００－ ７／５０