

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4110734号
(P4110734)

(45) 発行日 平成20年7月2日(2008.7.2)

(24) 登録日 平成20年4月18日(2008.4.18)

(51) Int.Cl. F I
H04L 12/56 (2006.01) H04L 12/56 230A

請求項の数 4 (全 24 頁)

(21) 出願番号	特願2000-358779 (P2000-358779)	(73) 特許権者	000000295 沖電気工業株式会社 東京都港区虎ノ門1丁目7番12号
(22) 出願日	平成12年11月27日(2000.11.27)	(74) 代理人	100090620 弁理士 工藤 宣幸
(65) 公開番号	特開2002-164921 (P2002-164921A)	(72) 発明者	横山 篤史 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内
(43) 公開日	平成14年6月7日(2002.6.7)	(72) 発明者	青柳 弘美 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内
審査請求日	平成18年1月31日(2006.1.31)	(72) 発明者	山口 順以 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

最終頁に続く

(54) 【発明の名称】 音声パケット通信の品質制御装置

(57) 【特許請求の範囲】

【請求項1】

品質非保証型のネットワークを介して音声パケットを伝送する音声パケット通信の品質制御装置において、

前記ネットワークを介して受信された音声パケットを一時的に蓄積して、受信した音声パケットの待ち行列を形成するためのバッファメモリと、

供給を受ける操作制御信号に応じて、当該待ち行列に対する操作を行う待ち行列操作手段と、

前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列が持つ音声的な特性を検査する系列検査手段と、

当該系列検査手段の検査結果に応じて前記操作制御信号を変更する操作制御手段とを備え、

前記操作制御手段は、

前記系列検査手段の検査結果を用いて、前記待ち行列上で分散するようにかつ音質に影響を与えないように操作位置を決定し、前記操作制御信号として操作位置指定信号を出力する操作位置決定部を備え、

前記待ち行列操作手段は、

供給を受ける当該操作位置指定信号に応じた前記待ち行列上の操作位置に存在する音声パケットを、当該待ち行列から削除する削除操作部および/または、

供給を受ける当該操作位置指定信号に応じた前記待ち行列上の操作位置に対し、所定の

音声情報を収容している補完音声パケットを挿入する挿入操作部を備えることを特徴とする音声パケット通信の品質制御装置。

【請求項 2】

品質非保証型のネットワークを介して音声パケットを伝送する音声パケット通信の品質制御装置において、

前記ネットワークを介して受信された音声パケットを一時的に蓄積して、受信した音声パケットの待ち行列を形成するためのバッファメモリと、

供給を受ける操作制御信号に応じて、当該待ち行列に対する操作を行う待ち行列操作手段と、

前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列が持つ音声的な特性を検査する系列検査手段と、

当該系列検査手段の検査結果に応じて前記操作制御信号を変更する操作制御手段と、
前記待ち行列の長さに対し少なくとも、上限側に設定する上限閾値を管理する閾値管理手段と、

当該待ち行列の長さ上限閾値との関係を監視する行列長監視手段とを備え、

前記操作制御手段は、

前記系列検査手段の検査結果を用いて、前記待ち行列上で分散するように操作位置を決定し、前記操作制御信号として操作位置指定信号を出力する操作位置決定部を備え、

前記系列検査手段は、

前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列を検査することにより、各音声パケットの復号時における重要度である復号重要度を検出する復号重要度検出部と、

当該復号重要度検出部が検出した復号重要度を、前記待ち行列を構成している各音声パケットに対応付けて一時的に記憶させる復号重要度記憶部とを備え、

前記待ち行列操作手段は、

前記行列長監視手段が前記待ち行列が上限閾値よりも長いことを検出した場合、重要度が低い復号重要度に対応付けられた音声パケットほど優先的に当該待ち行列から削除する優先削除操作部を備えることを特徴とする音声パケット通信の品質制御装置。

【請求項 3】

品質非保証型のネットワークを介して音声パケットを伝送する音声パケット通信の品質制御装置において、

前記ネットワークを介して受信された音声パケットを一時的に蓄積して、受信した音声パケットの待ち行列を形成するためのバッファメモリと、

供給を受ける操作制御信号に応じて、当該待ち行列に対する操作を行う待ち行列操作手段と、

前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列が持つ音声的な特性を検査する系列検査手段と、

当該系列検査手段の検査結果に応じて前記操作制御信号を変更する操作制御手段と、
前記待ち行列を構成する音声パケットの伝送方向に対応する音声受信経路上の音声信号および当該音声受信方向と反対の音声送信経路上の音声信号につき、有音無音判定を行って、音声受信経路上の音声信号と音声送信経路上の音声信号の双方が有音状態となるデュアルトーク期間の長さの伸縮傾向を検出するデュアルトーク期間伸縮傾向検出手段と、

当該待ち行列の長さに対し少なくとも、上限側に設定する上限閾値を管理する閾値管理手段と、

当該上限閾値を変化させる第 1 の上限閾値変更手段と、

当該待ち行列の長さ上限閾値との関係を監視する行列長監視手段とを備え、

前記操作制御手段は、

前記系列検査手段の検査結果を用いて、前記待ち行列上で分散するように操作位置を決定し、前記操作制御信号として操作位置指定信号を出力する操作位置決定部を備え、

当該デュアルトーク期間伸縮傾向検出手段が検出する伸縮傾向に応じて、前記第 1 の上

10

20

30

40

50

限閾値変更手段が、前記上限閾値を変化させることを特徴とする音声パケット通信の品質制御装置。

【請求項 4】

品質非保証型のネットワークを介して音声パケットを伝送する音声パケット通信の品質制御装置において、

前記ネットワークを介して受信された音声パケットを一時的に蓄積して、受信した音声パケットの待ち行列を形成するためのバッファメモリと、

供給を受ける操作制御信号に応じて、当該待ち行列に対する操作を行う待ち行列操作手段と、

前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列が持つ音声的な特性を検査する系列検査手段と、

当該系列検査手段の検査結果に応じて前記操作制御信号を変更する操作制御手段と、

前記待ち行列を構成する音声パケットの伝送方向に対応する音声受信経路上の音声信号および当該音声受信方向と反対の音声送信経路上の音声信号につき、音声パワーの検出を行って、単位時間あたりに音声受信経路上の音声パワーと音声送信経路上の音声パワーの大小関係が変化する回数の増減傾向を検出する増減傾向検出手段と、

前記待ち行列の長さに対し少なくとも、上限側に設定する上限閾値を管理する閾値管理手段と、

当該増減傾向検出手段が検出する増減傾向に応じて、上限閾値を変化させる第2の上限閾値変更手段とを備え、

前記操作制御手段は、

前記系列検査手段の検査結果を用いて、前記待ち行列上で分散するように操作位置を決定し、前記操作制御信号として操作位置指定信号を出力する操作位置決定部を備えることを特徴とする音声パケット通信の品質制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は音声パケット通信の品質制御装置に関し、例えば、インターネットなどのパケット網を利用して音声パケット通信を行う場合などに利用して好適なものである。

【0002】

【従来の技術】

現在、インターネットなどのパケット網にて音声信号をリアルタイム伝送するための技術が提案され、実際の装置として供されている。

【0003】

【発明が解決しようとする課題】

ところがインターネットは元々、リアルタイム伝送を必要としないデータ通信のために考案されたものであり、インターネット上のパケット伝送は品質保証されていない。このため、インターネット上では、パケットの欠落（パケット損失）、遅延、ジッタなどの復号音声の品質を劣化させる要因となる現象が発生し得る。

【0004】

したがって、このようなインターネットを電話などのリアルタイム性を要求される通信機能のために活用しようとする場合、伝送途切れを防止するためのバッファ装置を必要とする。

【0005】

このバッファ装置は、インターネットなどのネットワークから受信した音声パケット（この音声パケットは、符号化音声データ（多くの場合、当該符号化音声データは、非可逆圧縮符号化方式によって圧縮された音声データ）を収容しているパケット）を受信順にしたがって蓄積しておき、蓄積順に読み出す構成としておく。この場合、当該読み出しは必ず、当該符号化音声データを復号（解凍）する機能を持つ復号回路が要求する一定の復号単位時間毎に繰り返し実行されることになる。

10

20

30

40

50

【0006】

したがって当該バッファ装置を用いる場合、前記ジッタの影響などにより、音声パケットの受信装置への到着が一定時間以上遅れると、バッファ装置に対する音声パケットの蓄積は行われず、読み出しのみが進行する状態となり、読み出すべき音声パケットが枯渇してしまうことがある。

【0007】

このような枯渇が発生した場合でも前記復号回路に対しては前記復号単位時間間隔で音声パケットを供給しつづける必要があるため、そのようなケースでは、所定の音声データ（多くの場合は無音に近い微少な雑音を復号音声として発生する音声データ）を収容した補完パケットを挿入する技術が一般に利用されている。

10

【0008】

しかしながら当該補完パケットを挿入した場合、前記ジッタによって到着が遅れたパケットが後で届けられるために、バッファ装置中のパケットが徐々に増加して伝送遅延が時間とともに長くなっていくという問題がある。

【0009】

当該伝送遅延が長くなると、例えば、双方向の会話音声において発話した内容に対する応答が不自然に遅れ、通信の品質が低下してしまう。

【0010】

この対策としては、蓄積される音声パケットの数が所定数よりも多くなると当該バッファ装置中で最先に蓄積された（すなわち最先位置の）音声パケットを削除（廃棄）する操作を行うことが考えられる。

20

【0011】

また、音声パケットの到着が一定時間以上遅れ、読み出すべき音声パケットが枯渇した場合に前記補完パケットを挿入する位置も、当該最先位置にすることが考えられる。

【0012】

しかしながらこのように最先位置についてだけ補完パケットの挿入や音声パケットの削除を実行すれば、処理が簡略化できる利点がある反面、当該最先位置の状態（またはそれ以前に読み出された音声パケットの状態）しか監視できないため、結果的に、音声パケット系列上の特定の位置に連続して削除や挿入を行う可能性が高まる。

【0013】

削除すれば、復号時に必要とされる有効な音声データが失われ得るし、挿入すれば復号時に不要な音声データが混入されるので、どちらも復号音声出力の品質を劣化させ得る操作であるが、音声パケット系列上に連続して削除や挿入を行った場合には、その劣化の程度が著しいものとなる可能性が高い。

30

【0014】

【課題を解決するための手段】

かかる課題を解決するために、本発明では、品質非保証型のネットワークを介して音声パケットを伝送する音声パケット通信の品質制御装置において、（１）前記ネットワークを介して受信された音声パケットを一時的に蓄積して、受信した音声パケットの待ち行列を形成するためのバッファメモリと、（２）供給を受ける操作制御信号に応じて、当該待ち行列に対する操作を行う待ち行列操作手段と、（３）前記バッファメモリ中に蓄積されている待ち行列を構成する複数の音声パケットが収容している音声情報の系列が持つ音声的な特性を検査する系列検査手段と、（４）当該系列検査手段の検査結果に応じて前記操作制御信号を変更する操作制御手段とを備えることを特徴とする。

40

【0015】

【発明の実施の形態】

（Ａ）実施形態

以下では本発明にかかる音声パケット通信の品質制御装置をインターネット上に配置された音声通信装置に適用した場合を例に、実施形態について説明する。

【0016】

50

当該音声通信装置は、例えば、V o I P (Voice Over IP) に対応するもので、インターネットを介して双方向または一方向で会話音声やり取りされる。一例として、当該音声通信装置は、インターネット電話などであってよい。

【 0 0 1 7 】

(A - 1) 第 1 の実施形態の構成

本実施形態の音声通信装置 1 2 を含む音声通信システム 1 0 の主要部の構成例を図 1 に示す。

【 0 0 1 8 】

図 1 において、当該音声通信システム 1 0 は、音声通信装置 1 1 と、ネットワーク 1 5 と、当該音声通信装置 1 2 とを備えている。

10

【 0 0 1 9 】

図 1 に示した状態では、音声通信装置 1 1 は送信側として機能し、音声通信装置 1 2 は受信側として機能している。音声通信装置 1 1 は送信専用通信装置であり、音声通信装置 1 2 は受信専用通信装置であってもよいが、ともに送受信機能を搭載した送受信装置であってもよい。送受信装置である場合の構成や動作は、図 1 において音声通信装置 1 1 と 1 2 を相互に置き換えて複合すればよいだけであるので、以下では音声通信装置 1 1 が送信側として機能するとともに、音声通信装置 1 2 が受信側として機能する場合についてのみ説明する。

【 0 0 2 0 】

また、音声通信装置 1 1 と 1 2 とを接続しているネットワーク 1 5 は、例えば、インターネットなどの品質非保証型のパケット通信を行うネットワークであればどのようなネットワークであってもよいが、本実施形態ではインターネットであるものとする。

20

【 0 0 2 1 】

送信側である音声通信装置 1 1 は、マイクロフォン 1 3 と音声符号化器 1 4 とを備えている。

【 0 0 2 2 】

ここで、音声符号化器 1 4 は、集音されてマイクロフォン 1 3 から送出される音声データをインターネット 1 5 に送出するために符号化 (非可逆圧縮符号化) し、符号化音声データを一定の符号化単位時間ずつに分割して音声パケット P I に収容し、順次、インターネット 1 5 に送出する部分である。

30

【 0 0 2 3 】

非可逆圧縮の場合、データの欠落 (例えばビット単位の欠落) が生じる可能性があるが圧縮するとデータ量は例えば数十分の一から数百分の一程度 (可逆圧縮 (Lossless Compression) 方式の場合には例えば数分の一程度) になる。音声データは最終的にはパケット通信装置 1 2 のユーザなどの人間の聴覚を用いて感得されるものなので、厳密に圧縮前の状態に復元されず、データの一部が欠落したとしても、それが許容限度内であれば問題はない。したがって解凍後のデータの正確さよりむしろ、高い圧縮率によってデータのサイズを小さくし、実質的な伝送効率を高めるのに好都合な非可逆圧縮を行ったほうが、通信のリアルタイム性の向上などの点で有利である。

【 0 0 2 4 】

前記音声符号化器 1 4 の符号化単位時間 E T は通常、後述する音声復号器 1 7 における復号単位時間 D T に一致させるものである。例えば I T U - T 標準の G . 7 2 9 の場合、当該復号単位時間 D T は、 $D T = 10$ ミリ秒となり、G . 7 2 3 . 1 の場合は $D T = 30$ ミリ秒となるので、このうちのいずれかに設定するのが一般的である。

40

【 0 0 2 5 】

本実施形態では、 $D T = 10$ ミリ秒とする。この場合、前記音声パケット P I の長さは、実質的に当該 10 ミリ秒であるとみなすことができる。

【 0 0 2 6 】

音声通信装置 1 1 が送出した時系列な音声パケット P I をインターネット 1 5 を介して受信する音声通信装置 1 2 は、揺らぎ吸収バッファ装置 1 6 と、音声復号器 1 7 と、スピー

50

カ 18 と、補完パケット挿入部 19 と、パケット削除器 20 と、有音・無音判定器 21 とを備えている。

【0027】

このうち（揺らぎ吸収）バッファ装置 16 は、インターネット 15 にて発生する音声パケット P I の伝送遅延差（伝送遅延揺らぎ、すなわちジッタ）を吸収する機能を持つバッファメモリ 32 を主体とする装置で、その構成例は図 6 に示す通りである。

【0028】

（A - 1 - 1）揺らぎ吸収バッファ装置の内部構成

図 6 において、当該バッファ装置 16 は、キュー長検出部 30 と、走査読出部 31 と、当該バッファメモリ 32 とを備えている。

10

【0029】

このバッファメモリ 32 は基本的に F I F O（先入れ先出しタイプのメモリ）として機能するメモリで、インターネット 15 から受信された時系列な音声パケット P I は受信されたものから順番に当該バッファメモリ 32 に書き込まれ、書き込まれた順番に当該バッファメモリ 32 から読み出される。読み出された音声パケット P I は音声パケット P O として前記音声復号器 17 に供給される。

【0030】

当該読み出しは必ず、一定の前記復号単位時間毎に繰り返し実行されることになる反面、当該書き込みのほうはインターネット 15 上におけるパケットの欠落やジッタなどの影響を受けるため、一定の時間間隔で行われる保証はない。欠落やジッタの頻度や程度は、インターネット 15 のトラフィックの変化などに応じて時々刻々と変動するからである。

20

【0031】

受信された音声パケット P I の系列のうち、図 6 の状態で最先にバッファメモリ 32 に書き込まれた音声パケットが音声パケット P 1 であり、2 番目に書き込まれた音声パケットは音声パケット P 2 であり、3 番目に書き込まれた音声パケットは音声パケット P 3 であり、4 番目に書き込まれた音声パケットは音声パケット P 4 であり、...、100 番目に書き込まれた音声パケットは音声パケット P 100 であり、101 番目に書き込まれた音声パケットは音声パケット P 101 であり、最後に書き込まれた音声パケットは音声パケット P 102 である。

【0032】

すなわちこの状態では、バッファメモリ 32 内に 102 個の音声パケット P I が蓄積されており、キュー長（待ち行列の長さ）はパケット数で表現すると 102 パケット、復号音声出力の時間で表現すると 1.02（= 102 × 10 ミリ秒）秒ということになる。

30

【0033】

したがって図 6 上では、右側の音声パケットほど書き込みおよび読み出しの時刻が早く、左側の音声パケットほど書き込みおよび読み出しの時刻が遅くなる。なお、バッファメモリ 32 中で左端の空白部分は、音声パケットを書き込み得る状態にありながら現時点では音声パケットが書き込まれていないバッファ領域に対応している。

【0034】

本実施形態では、当該バッファメモリ 32 に上位、下位 2 つの閾値 T H と T L が設定されているが、下位閾値 T L のほうは発明が解決しようとする課題で述べた最先位置に相当する位置、すなわち、1 パケットと 0 パケットの中間のキュー長に相当する位置に設定されている。

40

【0035】

この下位閾値 T L を基準として、今回の読み出しでキュー長が 1 パケットでなくなり、それから復号単位時間後の次の読み出しでは、読み出すべきパケットが存在しない状態、すなわち（音声パケットの）枯渇が発生したことを検出するものである。

【0036】

上位閾値 T H のほうは、一例として、100 パケットと 99 パケットの中間のキュー長に相当する位置に設定されている。この上位閾値 T H の位置は、必要に応じて、上位方向ま

50

たは下位方向に静的に変更することが可能である。

【 0 0 3 7 】

バッファメモリ 3 2 におけるキュー長と 2 つの閾値 T H、 T L の関係を監視しているキュー長検出部 3 0 は、前記バッファメモリ 3 2 と、補完パケット挿入器 1 9 およびパケット削除器 2 0 のあいだにあって、キュー長が下位閾値 T L より短くなったときには、前記最先位置に対し当該補完パケット挿入器 1 9 に補完パケット P P を挿入させ、上位閾値 T H より長くなったときにはキュー上の適切な位置から、当該パケット削除器 2 0 に音声パケットの削除を行わせる部分である。

【 0 0 3 8 】

このため、当該キュー長検出部 3 0 は、バッファメモリ 3 2 から供給を受ける検出信号 D 1 を用いてキュー長を検出し、補完パケット挿入器 1 9 に供給する制御信号 C 2 を非能動状態から能動状態に切替えることによって当該補完パケット挿入器 1 9 に補完パケット P P の挿入を指示し、パケット削除器 2 0 に供給する制御信号 C 4 を非能動状態から能動状態に切替えることによって当該パケット削除器 2 0 に音声パケットの削除を指示する。

【 0 0 3 9 】

検出信号 D 1 の具体例としては、例えば、バッファメモリ 3 2 に対する音声パケット P I の書き込み数から音声パケット P I (P O) の読み出し数を減算することによって得られる差を用いることができる。

【 0 0 4 0 】

当該キュー長検出部 3 0 はまた、走査読出部 3 1 に供給する制御信号 C 3 を非能動状態から能動状態に切替えることによって、走査読み出し部 3 1 を機能させ、バッファメモリ 3 2 から走査信号 S C の読み出しを行う。

【 0 0 4 1 】

当該キュー長検出部 3 0 に接続された走査読出部 3 1 は、キュー長検出部 3 0 から供給される制御信号 C 3 が非能動状態から能動状態に切り替わると、バッファメモリ 3 2 に出力する制御信号 C 1 を非能動状態から能動状態に切替えて、その時点でキューを構成している音声パケットを次々と読み出し、読み出された音声パケットによって構成される走査信号 S C をパケット削除器 2 0 に供給する。

【 0 0 4 2 】

走査信号 S C を構成する音声パケット P I の読み出しは、前記復号単位時間に比べて十分に短い時間で行われる必要があり、一例としては、1 復号単位時間のあいだに 1 0 0 パケット程度の音声パケット P I を読み出すものであってよい。

【 0 0 4 3 】

もともと 1 0 ミリ秒 (すなわち、1 復号単位時間) に 1 つの音声パケット P I を読み出すのは、ユーザに聴取させる音声出力を発生するための音声復号器 1 7 の動作速度 (復号速度) に合わせるためであり、バッファメモリ 3 2 自体の性能としてはこれよりもはるかに高速な読み出しが可能である (例えば、当該バッファメモリ 3 2 として速度の遅い C M O S などを用いたとしても、1 0 0 ナノ秒のオーダーで読み出しが可能である) ので、このような走査信号 S C は、十分に実現可能である。

【 0 0 4 4 】

なお、キューを構成している音声パケット P I がバッファメモリ 3 2 から音声パケット P O として読み出された場合にはキューが処理され、読み出された音声パケット P I の分だけキュー長が短縮されるが、当該走査読み出し部 3 1 により走査信号 S C として読み出されてもキューは処理されず、キュー長は当該読み出しの前後で相違しない。

【 0 0 4 5 】

本実施形態の場合、上位閾値 T H の位置が 1 0 0 パケットと 9 9 パケットのあいだに設定されていることから、走査読み出し部 3 1 によって読み出される音声パケットは 1 0 0 パケットを超えることになるが、必要ならばこの数に制限を加えるようにしてもよい。

【 0 0 4 6 】

一例としては、最上位から 3 0 パケット (図 6 の例では、音声パケット P 1 0 2 から音声

10

20

30

40

50

パケット P 7 3 まで) だけを読み出して当該走査信号 S C としてもよい。

【 0 0 4 7 】

なお、図 6 では、バッファメモリ 3 2 に対して音声パケット P I を書き込む位置と、音声パケット P I を読み出す位置が相違しており、デュアルポートメモリを示しているようにもみえるが、図 6 はそのようなハードウェア的な構造を特定するものではない。

【 0 0 4 8 】

バッファメモリ 3 2 は F I F O として機能することができれば、どのようなハードウェア的な構造を備えていてもかまわない。例えば、図 6 の状態において隣接している領域(例えば、音声パケット P 1 が記憶されている領域と、音声パケット P 2 が記憶されている領域)は、必ずしも実際のバッファメモリ(3 2)上で物理的に隣接している必要はない。一般的にキューは、リスト構造を用いて論理的に実現されるものだからである。

10

【 0 0 4 9 】

また、音声パケット P I を読み出しただけではそのデータ内容はバッファメモリ 3 2 に残っているので、走査信号 S C の読み出し後は、バッファメモリ 3 2 に対して何も処理を施す必要はない。この観点でキューを構成していた音声パケット P I を音声パケット P O として読み出す場合の処理を説明すると、やはり音声パケット P I を読み出しただけではバッファメモリ 3 2 内に読み出した音声パケット P I のデータは残っているが、P O としての当該読み出し後、読み出した音声パケット P I が格納されていたバッファメモリ 3 2 上の領域は、以降に書き込まれる音声パケット P I によって占有され得る状態(すなわち書き込み可能状態)になるということができる。

20

【 0 0 5 0 】

以上のようなバッファ装置 1 6 において、前記枯渇が発生する可能性を低減するためにはキューは長いほどよいが、キューが長すぎると、実質的に前記伝送遅延が長くなり、例えば、双方向の会話音声において発話した内容に対する応答(この応答は、音声通信装置 1 1 から音声パケット P I によってもたらされる)が不自然に遅れる可能性が高まるので、当該バッファ装置 1 6 は、前記補完パケット挿入器 1 9 やパケット削除器 2 0 とも連携して、常に一定量のパケット(一定長のキュー)を保持するようにバッファメモリ 3 2 を制御する。

【 0 0 5 1 】

一方、キュー長検出部 3 0 から制御信号 C 4 を受け取るとともに走査読み出し部 3 1 から走査信号 S C を受け取る図 1 のパケット削除器 2 0 は、当該走査信号 S C を有音・無音判定器 2 1 に供給して、その判定結果 D C を受け取り、判定結果 D C に基づいてキュー上で削除する音声パケットを決定し、当該決定に応じた制御信号 C 5 を出力する部分である。

30

【 0 0 5 2 】

すなわちパケット削除器 2 0 は、バッファメモリ 3 2 中で保持される音声パケット P I が、上位閾値 T H よりも多くなったとき(すなわち、前記制御信号 C 4 が能動状態になったとき)に、バッファメモリ 3 2 の各々のパケットを有音・無音判定器 2 1 に問い合わせることで、無音と判定された音声パケットを削除する。

【 0 0 5 3 】

このとき、パケット削除器 2 0 は、音声復号器 1 7 に渡す直前の音声パケット(すなわち、前記最先位置にある音声パケット(P 1))を評価および削除するのみにとどまらず、前記走査信号 S C を用いてバッファメモリ 3 2 に蓄積されている複数の音声パケットを走査し、各々のパケットの前後のパケットの有音・無音判定結果 D C や、バッファメモリ 3 2 中の無音パケット比率に基づいて、削除パケットを決定する。

40

【 0 0 5 4 】

なお、一般に音声をパケット単位で復号する装置では、前記復号単位時間に対応する一定時間間隔でパケットを復号処理することが必要であるが、本実施形態を適用することにより、パケットの有音・無音判定および、パケットの削除を先行して処理することが可能となる。すなわち、処理装置の負荷が低いときに、これらを処理することが可能であるため、処理装置の稼働率を向上させることができる。

50

【 0 0 5 5 】

さらに、削除する音声パケット P I の決定に際して当該パケット削除器 2 0 は、削除される音声パケット P I ができるだけキュー上で連続せず、分散するように、またできるだけ無音と判定された音声パケットを削除しながら必要な数（削除数）の音声パケットを削除するように制御する。

【 0 0 5 6 】

ただし状況によってはこれらの条件のすべてを満足することが不可能なケースも起こり得る。例えば、キュー長が上位閾値 T H を超えたためにバッファ装置 1 6 から供給される制御信号 C 4 が能動状態になったが、判定結果 D C は音声パケット P 1 ~ P 1 0 2 のすべてが有音であることを示しているようなケースである。

10

【 0 0 5 7 】

このような場合には、より無音らしいパケットを無音とみなして処理を進めるようにするとよい。その場合でも、無音らしいために削除する音声パケットがキュー上で連続しないように制御することが必要である。

【 0 0 5 8 】

このような削除を行った結果として、キュー上のある位置には後述する有音らしさの低い（すなわち無音らしい）パケットが残留し、キュー上のある位置からは残留した有音らしさの低いパケットに比べると有音らしさの高いパケットが削除されることも起こり得るが、キュー全体としてみたとき、復号時の音声出力の品質は、キュー上で連続するパケットの削除を許容する場合に比べて向上する可能性が高い。

20

【 0 0 5 9 】

なお、ここで、削除とは、当該削除の対象となる音声パケット P I が格納されているバッファメモリ 3 2 上の領域を、前記書き込み可能状態に移行させる操作をいう。したがって削除時の前記制御信号 C 5 は、一般的には、バッファメモリ 3 2 上の多数のメモリアドレス（1つのメモリアドレスで指定される領域に1つの音声パケット P I が書き込まれている）のうち、1または複数のメモリアドレスを指定して、当該メモリアドレスで指定されるメモリ領域を書き込み可能状態に移行させることになる。

【 0 0 6 0 】

パケット削除器 2 0 から走査信号 S C を受け取る有音・無音判定器 2 1 は、パケット削除器 2 0 から問い合わせられたパケット（当該走査信号 S C に収容されているパケット）を一度復号し、あるいは音声パワー情報を取り出すことで、当該パケットが有音状態に属するものか、無音状態に属するものかを判定する部分である。

30

【 0 0 6 1 】

一般には音声パワー履歴から背景雑音強度を推定し、そこから推定 S / N（信号対雑音）比を求め、この推定 S / N 比を用いて有音・無音判定を行うものである。

【 0 0 6 2 】

また、バッファ装置 1 6 から制御信号 C 2 を受け取る補完パケット挿入器 1 9 は、当該制御信号 C 2 が能動状態になると、無音に近い微少な雑音を復号音声として発生する音声データを収容している所定の補完パケット P P を生成し、出力する部分である。

【 0 0 6 3 】

本実施形態において当該補完パケット挿入器 1 9 が出力した補完パケット P P は、固定的に、前記最先位置に挿入されるものとする。

40

【 0 0 6 4 】

また、バッファ装置 1 6 から前記復号単位時間ごとに音声パケット P O を受け取る音声復号器 1 7 は、当該音声パケット P O を所定の手順にしたがって復号し、復号音声出力をスピーカ 1 8 から発生させる部分である。

【 0 0 6 5 】

以下、上記のような構成を有する本実施形態の動作について説明する。

【 0 0 6 6 】

(A - 2) 第 1 の実施形態の動作

50

前記キュー長検出部30は、前記検出信号D1を用いて常にリアルタイムでバッファメモリ32に蓄積されている音声パケットPIによって構成されているキューのキュー長を監視している。

【0067】

そして、当該キュー長が上位閾値THよりも短く、下位閾値TLよりも長いときにはキュー長検出部30から出力される制御信号C2～C4はすべて非能動状態を維持し、走査読み出し部31から出力される制御信号C1も非能動状態を維持するので、図1上の補完パケット挿入器19、パケット削除器20、有音・無音判定器21は動作しない。

【0068】

したがって、図6に示す音声パケットP1～P99がバッファメモリ32に蓄積されているときも、補完パケット挿入器19、パケット削除器20、有音・無音判定器21は動作しない。

【0069】

音声パケットP1～P99が蓄積されている状態において、1復号単位時間が経過し、最初に蓄積された音声パケットP1が音声パケットPOとして読み出されたあとで音声パケットP100がインターネット15から受信され、バッファメモリ32に蓄積されると、当該バッファメモリ32内には音声パケットP2～P100によってキュー長が99パケットのキューが存在することになるが、この状態でも、補完パケット挿入器19、パケット削除器20、有音・無音判定器21は動作しない。

【0070】

ところが、音声パケットP1～P99が蓄積されている状態で、1復号単位時間が経過して最初に蓄積された音声パケットP1が読み出されるまえに、例えば、図6に示すように、音声パケットP100～P102がインターネット15から受信されバッファメモリ32に蓄積されたとすると、キュー長は102パケットとなり、上位閾値THを超える。このように、本来ならば1音声パケット(ここではP100)だけが受信されるはずの期間内に複数の音声パケット(P100～P102)が受信される状況は、インターネット15上で発生するジッタによって起こり得る。

【0071】

キュー長が上位閾値THを超えたことを検出信号D1に基づいて検出したキュー長検出部30は、制御信号C3とC4を非能動状態から能動状態に切替える。

【0072】

当該制御信号C4が能動状態に切り替わったことを検出したパケット削除器20は走査信号SCの供給を待つ。

【0073】

また、制御信号C3が能動状態に切り替わったことを検出した走査読み出し部31は、制御信号C1を非能動状態から能動状態に切替えて、現時点でキューを構成しているすべての音声パケットP1～P102を走査信号SCとして読み出し、当該走査信号SCをパケット削除器20に供給する。

【0074】

走査信号SCの供給を受けたパケット削除器20は当該走査信号SCを有音・無音判定器21に供給し、当該有音・無音判定器21から判定結果DCを受け取る。

【0075】

判定結果DCを受け取ったパケット削除器20は、上述した条件をできるだけ満足するように削除を行う。すなわち、削除する音声パケットPIができるだけキュー上で連続せず、分散する内容を持ち制御信号C5であって、またできるだけ無音と判定された音声パケットを削除するような内容を持つ制御信号C5を出力する。

【0076】

例えば、音声パケットP1～P102のうち、音声パケットP1～P5およびP98～P102が無音であるとの判定結果DCを得た場合、削除する必要がある音声パケットの数(削除数)を3とすると、例えば音声パケットP4、P99、P101を削除し、音声パ

10

20

30

40

50

ケット P 1 ~ P 1 0 2 によって構成されているキュー長 1 0 2 パケットのキューを、音声パケット P 1、P 2、P 3、P 5 ~ P 9 8、P 1 0 0、および P 1 0 2 によって構成されるキュー長 9 9 パケットのキューに変換する。

【 0 0 7 7 】

なお、削除数をどのように選定するかも問題であるが、削除自体が復号音声出力の劣化要因にもなることを考慮すると、必要最小限（キュー長が上位閾値 T H を超えなくなる最少の削除数）にとどめたほうがよいと考えられる。

【 0 0 7 8 】

ただし削除数を小さくしたために、削除によっていったん上位閾値 T H を下回ったキュー長がふたたび上位閾値 T H を超える状況が頻発し、走査信号 S C の読み出しにともなうパケット削除器 2 0 や有音・無音判定器 2 1 の処理に過大な負荷がかかる場合には、削除数を多くしてこの負荷を軽減するようにしてもよい。

【 0 0 7 9 】

本実施形態では、前記最先位置の音声パケット（図 6 の音声パケット P 1 に相当する音声パケット）だけでなくキューを構成している多数の音声パケットから構成される音声パケット群を分析し、当該分析結果に基づいて削除すべき音声パケットを決定することができるため、削除する音声パケットの位置をキュー上に分散させることが可能になる。

【 0 0 8 0 】

一方、キューを構成している音声パケットが図 6 の音声パケット P 1 だけである場合には、今回の読み出しでキュー長が 1 パケットでなくなり、それから復号単位時間後の次の読み出しでは、読み出すべきパケットが存在しない。すなわち、音声パケットの枯渇が発生している。

【 0 0 8 1 】

検出信号 D 1 を用いて当該枯渇を検出したキュー長検出部 3 0 は、制御信号 C 2 を非能動状態から能動状態に切替えて補完パケット挿入器 1 9 に補完パケット P P を生成させ、出力させる。

【 0 0 8 2 】

補完パケット挿入器 1 9 が当該補完パケット P P を挿入する位置は、当該音声パケット P 1 が存在する最先位置である。

【 0 0 8 3 】

以上の動作の説明では、キュー長が上位閾値 T H を超えたときに当該キューを構成している 1 0 0 パケット程度のすべての音声パケットを走査信号 S C として読み出して一括して処理するようにしたが、もしもパケット削除器 2 0 や有音・無音判定器 2 1 における処理時間を確保するために必要ならば、1 0 0 パケット程度の音声パケットを分割して処理するようにしてもよい。

【 0 0 8 4 】

例えば、上位閾値 T H より下位（例えばキュー長が 6 0 パケット程度の位置に）に予備的な閾値 S T を設定しておき、当該予備的な閾値 S T をキュー長が超えた段階で、その時点のキューを構成している音声パケットの全部または一部を走査信号 S C として読み出してその判定結果 D C を得ておき、次いでキュー長が上位閾値 T H を超えたときには、判定結果 D C を得ていない音声パケットについてのみ、判定結果 D C を得て、パケット削除器 2 0 の処理を実行するようにしてもよい。

【 0 0 8 5 】

また、このような予備的な閾値 S T などを設定せずに、キュー長にかかわらず、処理能力に余裕のあるときに有音・無音判定を行い、判定結果 D C を蓄積しておくようにしてもよい。

【 0 0 8 6 】

ただしキュー長が上位閾値 T H を超えなければ、音声パケットの削除は行われないうえ、求めた判定結果 D C は対応する音声パケットが P O として読み出されたときに存在価値を失い、無駄になってしまうので、処理効率の観点では、キュー長が実際に上位閾値 T H を

10

20

30

40

50

超えたか、あるいは超える可能性が十分に高まり、なおかつ処理能力に余裕のあるときに、有音・無音判定を行うようにするのがよいと考えられる。

【0087】

(A-3) 第1の実施形態の効果

本実施形態によれば、削除する音声パケットの位置をキュー上に分散させることが可能になるので、特定箇所(例えば前記最先位置)の無音パケットを連続して削除する場合に比べて、より自然で、高品質な復号音声を得ることが可能になる。

【0088】

また、最先位置にある音声パケットのみに着目して無音パケットを選択すると、後続の音声パケットとの関係が不明であるため会話の語頭を誤って無音と判定してしまう危険性が高いが、本実施形態では最先位置以降の時間位置にある音声パケットをも分析した上で、削除すべき音声パケットを選択するので、語頭を容易に発見でき、キュー上で有音と判定された音声パケットの直前に位置している音声パケットは無音と判定されていても削除しない等の処理を行うことにより、語頭を誤って削除することを防止できる。

【0089】

さらに、前記最先位置で削除を行う場合、削除しなければ当該最先位置にある音声パケットをただちに復号しなければならぬため、短時間のうちに有音・無音判定、削除するかどうかの判断、復号処理などを集中して実行することを求められ、時間的に集中して負荷が発生することが避けられなかったが、本実施形態では、音声通信装置内での処理を時間的にずらして分散した上で実行することが容易であるので、優先度の高い音声復号器(17)の処理などが行われていない低負荷時間に、パケット削除器(20)や有音・無音判定器(21)の処理を実行することができ、時間的な負荷の分散を効率的に実現することが可能である。

【0090】

(B) 第2の実施形態

以下では、本実施形態が第1の実施形態と相違する点についてのみ説明する。

【0091】

(B-1) 第2の実施形態の構成および動作

図2は、本実施形態の揺らぎ吸収バッファ装置46の内部構成例を示している。

【0092】

このバッファ装置46は、第1の実施形態のバッファ装置16に対応する部分であり、基本的な機能は当該バッファ装置16と同じである。

【0093】

当該バッファ装置46を搭載した受信側の音声通信装置42を含む音声通信システム40の全体構成は、図7に示す通りである。

【0094】

すなわち、本実施形態の音声通信システム40の構成は、図1に示す第1の実施形態の音声通信システム10と基本的に同じである。

【0095】

図7において、本実施形態の音声通信システム40は、音声通信装置11と、インターネット15と、音声通信装置42とを備えている。

【0096】

図7中で図1と同じ符号を付与した各構成部分および各信号の機能は、図1と同じである。

【0097】

すなわち構成部分40, 42, 46, 50, 51以外の各構成部分および、信号DC1, SC1, SC2以外の各信号の機能は、本実施形態でも第1の実施形態と同じである。

【0098】

本実施形態の有音・無音判定部51は、図2に示す重要度設定部47から走査信号SC1を受け取ったときに2値的な有音・無音の判定結果DCではなく、3段階の「有音らしさ

10

20

30

40

50

」を示す判定結果DC1として、重要度を返す。ここで、有音らしさとは、有音・無音判定器51の処理の結果として得られる値に対応し、ある音声パッケージが収容している音声データが有音である蓋然性の高さを示している。

【0099】

ただし有音・無音判定器51は、搭載しているプログラムにしたがってこの「有音らしさ」を算出するので、その算出結果は実際に復号音声出力を人間の聴覚で感得したときに得られる結論と完全に一致している保証はない。

【0100】

また、前記「無音らしさ」とは、この「有音らしさ」と対立する概念で、ある音声パッケージが収容している音声データが無音である蓋然性の高さを示している。

10

【0101】

そして本実施形態のパケット削除器50は、キュー長が前記上位閾値THを超えた場合などの削除の必要性が発生したときに、バッファ装置46から、走査信号SC2として当該重要度を受け取り、当該重要度に基づいて削除する音声パッケージを決定し、削除を実行する部分である。

【0102】

したがって本実施形態のパケット削除器50と有音・無音判定器51とのあいだには、第1の実施形態の信号SCやDCのように、やり取りされる信号は存在しない。

【0103】

図2において、本実施形態のバッファ装置46は、キュー長検出部30と、走査読み出し部31と、バッファメモリ32Aと、重要度設定部47とを備えている。

20

【0104】

図2においても、図6と同一の符号を付した各構成部分30、31、TH、TL、および各信号PI、PO、C1~C5、D1の機能は図6と同じである。

【0105】

したがって本実施形態と第1の実施形態の相違は、主としてバッファメモリ32Aや重要度設定部47に関連する部分にかざられる。

【0106】

このうちバッファメモリ32Aは基本的に前記バッファメモリ32と同じ機能を備えたFIFOメモリであるが、キューを構成している各音声パッケージ(例えば、P1)に対応付けて、重要度設定部47から供給される重要度M1を格納する機能を備えている。各音声パッケージと各重要度との対応付けは、ハードウェア的に行われてもソフトウェア的に行われてもかまわない。

30

【0107】

本実施形態では、バッファメモリ32A中のキューを構成している音声パッケージの削除に先立って、当該重要度M1の設定が行われることを要する。

【0108】

図2中で、各音声パッケージP1~P102の内部に点線で示した枠がこの重要度を格納する部分であり、枠内の1~3の数値が重要度を示している。

【0109】

本実施形態では重要度は3段階で示し、数値が大きい重要度ほど対応する音声パッケージが収容している音声データの「有音らしさ」が高いことを示す。この重要度は、重要度設定部47が出力する重要度設定信号M1に応じてバッファメモリ32A内に設定された(書き込まれた)ものである。

40

【0110】

重要度設定部47は、重要度を設定する重要度設定タイミングがおとずれたことを、走査読み出し部31から走査信号SC1の供給を受けることによって検出する。当該走査信号SC1は、第1の実施形態の走査信号SCと同様にその時点でキューを構成している各音声パッケージによって構成される信号である。

【0111】

50

当該走査信号 S C 1 は重要度設定部 4 7 から有音・無音判定器 5 1 に供給され、有音・無音判定器 5 1 は、各音声パケットについての前記 3 段階の判定結果 D C 1 を当該重要度設定部 4 7 に返す。

【 0 1 1 2 】

当該判定結果 D C 1 を受け取った重要度設定部 4 7 は、当該判定結果 D C 1 に対応する前記重要度設定信号 M 1 を出力して、重要度 M 1 の設定を行う。

【 0 1 1 3 】

このあと、音声パケットの削除を行う場合には、走査読み出し部 3 1 が走査信号 S C 2 として各音声パケット（図 2 の例では、P 1 ~ P 1 0 2）に対応付けられている重要度 M 1 を読み出し、パケット削除器 5 0 に供給する。

10

【 0 1 1 4 】

パケット削除器 5 0 は当該重要度 M 1 をもとに削除を実行するが、一般に、パケット蓄積数が多くキュー長が上位閾値 T H よりも長いときには、パケット重要度が比較的高い音声パケットをも削除し、逆にキュー長が上位閾値 T H よりも短い場合には、音声パケットを一切削除しない。

【 0 1 1 5 】

また、当該パケット削除器 5 0 は、削除される音声パケットの位置がキュー上でできるだけ連続せず、分散するように、そして、できるだけ重要度の低い音声パケットが削除されるように、前記制御信号 C 5 を出力する。

【 0 1 1 6 】

20

本実施形態において、削除のタイミングは第 1 の実施形態で述べたものと同じであってよいが、重要度設定のタイミングは、削除と同じ（すなわち、例えば、キュー長が上位閾値 T H を超えたときに重要度の設定を行い、当該設定後ただちに削除を行う）であるか、または、削除より早期のいずれかの時点で実行することになる。

【 0 1 1 7 】

なお、図 2 中で、走査読み出し部 3 1 から出力される制御信号 C 1 0 は、当該走査信号 S C 1 を読み出すための信号であり、制御信号 C 1 は走査信号 S C 2 を読み出すための信号である。

【 0 1 1 8 】

（ B - 2 ）第 2 の実施形態の効果

30

本実施形態によれば、第 1 の実施形態の効果と同等な効果を得ることができる。

【 0 1 1 9 】

加えて、本実施形態では、3 段階の重要度を利用して削除する音声パケットを決定するので、削除による音質劣化の発生頻度と程度を低減することができ、有音、無音の 2 段階しか判断基準のない第 1 の実施形態に比べると、削除後に得られる復号音声出力の品質が高い。

【 0 1 2 0 】

また、例えば、パケット蓄積数が上位閾値に比べてかなり多いときには重要度の比較的高いパケット（例えば重要度 2 のパケット）も削除するが、わずかに多いときには重要度の低いパケットだけを削除する等、パケット蓄積数が上位閾値よりもどれくらい多いかに応じて、削除するパケットの重要度を変更することもできるので、柔軟な制御が可能になる。

40

【 0 1 2 1 】

（ C ）第 3 の実施形態

以下では、本実施形態が第 1 の実施形態と相違する点についてのみ説明する。

【 0 1 2 2 】

この相違点は、第 1 の実施形態では固定されていた前記上位閾値（ T H ）を動的に変更し得る点にかぎられる。

【 0 1 2 3 】

（ C - 1 ）第 3 の実施形態の構成および動作

50

本実施形態の音声通信システム 60 の主要部の構成例を図 9 に示す。

【0124】

図 9 において、図 1 と同じ符号を付した各構成部分および各信号の機能は、図 1 と同じである。

【0125】

したがって図 9 において、本実施形態の音声通信装置 62 が備える構成のうち前記音声通信装置 12 と相違するのは、マイクロフォン 63 と、音声符号化器 67 と、デュアルトーク検出器 64 を有している点と、揺らぎ吸収バッファ装置 66 の内部構成だけである。

【0126】

また、本実施形態の音声通信装置 61 は受信回路部 65 を備えている点が第 1 の実施形態の音声通信装置 11 と相違する。

10

【0127】

当該受信回路部 65 は、音声通信装置 62 における補完パケット挿入器 19 と、揺らぎ吸収バッファ装置 66 と、音声復号器 17 と、パケット削除器 20 と、有音・無音判定器 21 と同様な構成部分を内蔵する部分であってよい。

【0128】

音声通信装置 62 の構成部分のうちマイクロフォン 63 は前記マイクロフォン 13 に対応し、音声符号化器 67 は前記音声符号化器 14 に対応する。

【0129】

すなわち本実施形態の場合、音声通信装置 62 内において送信用と受信用の 2 つの経路が同時に有音状態となるデュアルトーク状態の期間が伸長しているか短縮しているかを検出することによって上位閾値 TH1 を変更するので、音声通信装置 61, 62 はともに、送信専用または受信専用の通信装置ではなく、送受信装置であることを要する。

20

【0130】

上述したように、前記枯渇が発生する可能性を低減するためにはキューは長いほどよいが、キューが長すぎると、実質的に前記伝送遅延が長くなり、例えば、双方向の会話音声において発話した内容に対する応答が不自然に遅れる可能性が高まるので、当該バッファ装置 66 は、前記補完パケット挿入器 19 やパケット削除器 20 とも連携して、常に一定量のパケット（一定長のキュー）を保持するようにバッファメモリ 32 を制御する。

【0131】

したがって第 1 の実施形態のバッファ装置 16 では、キュー長が上位閾値 TH を超えると、パケット削除器 20 から供給される制御信号 C5 を用いて音声パケットの削除を実行してキュー長を削減することになる。

30

【0132】

しかしながらキュー長が上位閾値 TH を超えたときに削除を行うということは、換言すると、上位閾値 TH 程度の遅延は許容しているということができる。

【0133】

ところが、電話などで双方向の会話音声やり取りされる状況においては、会話のパターンによって最適な上位閾値 TH の値（すなわち、許容される遅延）が相違することが判明している。

40

【0134】

本実施形態はこの点に着目してなされたもので、会話のパターンによって、前記上位閾値 TH に対応する上位閾値 TH1 の値を動的に変更することを特徴とする。

【0135】

図 9 の構成を有する本実施形態の音声通信システム 60 において、音声通信装置 61 側のユーザ（話者）が発話した音声は音声通信装置 62 側では、音声信号 DV に対応する音声出力としてスピーカ 18 から出力される。

【0136】

反対に、音声通信装置 62 側のユーザ（話者）が発話した音声は音声信号 EV として音声符号化器 67 に入力され、音声パケット PI1 に収容された上でネットワーク 15 を介し

50

て音声通信装置 6 1 側のユーザに聴取される。

【 0 1 3 7 】

音声通信装置 6 2 中のデュアルトーク検出器 6 4 は、音声符号化器の入力である音声信号 E V と音声復号器 1 7 の出力である音声信号 D V を同時に監視し、二つの経路（2 つの音声信号 E V と D V ）が同時に有音状態である時間（デュアルトーク時間）を累算して、デュアルトーク時間が長くなる傾向にあるか、短くなる傾向にあるかを検出して当該検出結果に対応する制御信号 C 1 1 を出力する部分である。

【 0 1 3 8 】

具体的には、例えば、過去 1 0 秒間の間のデュアルトーク時間を、1 秒間隔で計測し、デュアルトーク時間が前計測時（1 秒前）よりも長くなっている場合には、上位閾値 T H 1 を 1 パケット分だけ小さく変更する制御信号 C 1 1 を出力する。

10

【 0 1 3 9 】

逆に、デュアルトーク時間が前計測時よりも短くなっている場合には、出力する上位閾値 T H 1 を 1 パケット分だけ大きく変更する制御信号 C 1 1 を出力する。

【 0 1 4 0 】

当該制御信号 C 1 1 を受け取るのは、図 8 に示すバッファ装置 6 6 内のキュー長検出部 3 0 A である。キュー長検出部 3 0 A は当該制御信号 C 1 1 に応じて上位閾値 T H 1 を変更した上で、キュー長が変更後の上位閾値 T H 1 と比べて長いか、短いかを検査し、長い場合には前記制御信号 C 3 を非能動状態から能動状態に切替える。

【 0 1 4 1 】

20

例えば、図 8 に示した状態では、上位閾値 T H 1 は 1 0 0 パケットと 9 9 パケットの中間のキュー長に相当する位置に設定されているが、これを 1 パケット分だけ大きく変更すると、当該上位閾値 T H 1 は 1 0 1 パケットと 1 0 0 パケットの中間のキュー長に相当する位置に変更され、以降はこの変更後の上位閾値 T H 1 を基準として、キュー長と上位閾値 T H 1 の関係が検査され、1 0 1 番目の音声パケット P 1 0 1 が蓄積されたときにはじめて、パケットの削除が行われることになる。

【 0 1 4 2 】

(C - 2) 第 3 の実施形態の効果

本実施形態によれば、第 1 の実施形態と同等の効果を得ることができる。

【 0 1 4 3 】

30

加えて、本実施形態では、会話パターンに適応して、固定遅延（上位閾値（T H 1 ））を動的に変更することができるので、ユーザが不快感を感じやすいデュアルトーク状態が頻発しているときには、遅延を小さくして通信の品質を向上することが可能になる。

【 0 1 4 4 】

反対に、例えば、音声ガイダンスが流れているような状態（非デュアルトーク状態）では、固定遅延を大きくすることで、バッファ枯渇による音声の途切れが発生しにくくなるようにすることができる。

【 0 1 4 5 】

(D) 第 4 の実施形態

以下では、本実施形態が第 1 、第 3 の実施形態と相違する点についてのみ説明する。

40

【 0 1 4 6 】

(D - 1) 第 4 の実施形態の構成および動作

本実施形態の音声通信システム 7 0 の主要部の構成例を図 4 に示す。

【 0 1 4 7 】

図 4 において、図 9 と同じ符号を付与した各構成部分および各信号の機能は図 9 と同じである。

【 0 1 4 8 】

すなわち、本実施形態の音声通信システム 7 0 は、第 3 の実施形態の音声通信システム 6 0 の音声通信装置 6 2 内のデュアルトーク検出器 6 4 を、パワー変動差分演算器 7 1 と 0 交差カウンタ 7 2 で置換した構成を備えている。

50

【 0 1 4 9 】

このような構成を用いて、第 3 の実施形態ではデュアルトーク期間の伸縮傾向に対応した会話パターンの分析結果に応じて制御信号 C 1 1 を変更したのに対し、本実施形態では、会話の交代時間間隔に対応した会話パターンの分析結果に応じて制御信号 C 1 1 を変更する。

【 0 1 5 0 】

本実施形態のパワー変動差分演算器 7 1 は、符号化器側の音声信号 E V と復号器側の音声信号 D V について過去一定時間の音声パワーを求め、そのパワー差を出力するものである。

【 0 1 5 1 】

具体的には、それぞれの側において、例えば、過去 1 0 0 ミリ秒間の平均パワーを測定し、その差を出力する。符号化器側のパワーが大きいときは正の値を出力し、そうでないときは負の値を出力する。

【 0 1 5 2 】

また、前記 0 交差カウンタ 7 2 は、演算器 7 1 の出力が 0 を交差するたびに、その回数を累積する部分である。ただし、入力において 0 近傍に閾値 C L を設けてあり、小さな雑音などの、ある値よりも小さな振幅の入力変動は無視して、安定的に動作する。

【 0 1 5 3 】

当該 0 交差カウンタ 7 2 は、例えば、1 秒間隔でカウンタの値（累積した回数）に対応する制御信号 C 1 1 を出力し、当該制御信号 C 1 1 に応じて上位閾値 T H 2（前記上位閾値 T H 1 に対応する）を動的に変更することができる。そして 0 交差カウンタ 7 2 のカウンタ値は、当該制御信号 C 1 1 の出力後に 0 に初期化される。

【 0 1 5 4 】

これにより、0 交差カウンタ 7 2 のカウンタの値が、前計測時（1 秒前）のカウンタ累積値よりもより大きくなっている場合には、上位閾値 T H 2 を 1 パケット分だけ小さくするような制御信号 C 1 1 を出力し、逆に、前計測時よりも小さくなっている場合には、上位閾値 T H 2 を 1 パケット分だけ大きくするような制御信号 C 1 1 を出力する。

【 0 1 5 5 】

（ D - 2 ）第 4 の実施形態の効果

本実施形態によれば、第 1 の実施形態の効果と同等な効果を得ることができる。

【 0 1 5 6 】

加えて、本実施形態では、会話の交代時間間隔に対応した会話パターンの分析結果に対応して、固定遅延を動的に変更することが可能になる。

【 0 1 5 7 】

これにより、ユーザが不快感を感じやすい会話パターンでは固定遅延を小さくし、逆に、音声ガイダンスが流れているような状態では、固定遅延を大きくすることで、バッファ枯渇による音声の途切れを発生しにくくすることができる。

【 0 1 5 8 】

（ E ）第 5 の実施形態

以下では、本実施形態が第 1 の実施形態と相違する点についてのみ説明する。

【 0 1 5 9 】

（ E - 1 ）第 5 の実施形態の構成および動作

本実施形態の音声通信システム 8 0 の全体構成は図 5 に示し、バッファ装置 8 6 の内部構成例は図 3 に示す。

【 0 1 6 0 】

図 5 において、図 1 と同じ符号を付した各構成部分および各信号の機能は、図 1 と同じである。したがって図 5 上で本実施形態が第 1 の実施形態と相違する点は、バッファ装置 8 6、補完パケット挿入器 1 9 A、有音・無音判定器 2 1 A に関連する部分に限られる。

【 0 1 6 1 】

またこの図 5 上では、補完パケット挿入部 1 9 A に有音・無音判定器 2 1 A が接続されて

10

20

30

40

50

いて、バッファ装置 8 6 から補完パケット挿入部 1 9 B に対し走査信号 S C 3 が供給され、有音・無音判定器 2 1 A と当該補完パケット挿入器 1 9 A とのあいだで、走査信号 S C 3 と判定結果 D C 3 がやり取りされる。

【 0 1 6 2 】

次に、図 3 においても、図 6 と同じ符号を付した各構成部分および各信号の機能は図 6 と同じである。

【 0 1 6 3 】

したがって図 3 上で本実施形態が第 1 の実施形態と相違する点は、下位閾値 T H 1 と、キュー長検出部 8 2 と、走査読み出し部 8 3 に関連する部分にかざられる。

【 0 1 6 4 】

また、図 3 には本実施形態の特徴である下位閾値 T L 1 に関連する構成部分および信号だけを示し、上位閾値 T H に関連する構成部分および信号については、第 1 の実施形態と同じであるので、図示を省略している。

【 0 1 6 5 】

本実施形態と第 1 の実施形態の相違は、本実施形態が下位閾値 T L 1 を前記最先位置よりも上位の位置に設定したことから生じたものである。

【 0 1 6 6 】

上述したように、通常、前記補完パケット P P は、前記枯渇が発生した場合でも前記音声復号器 1 7 に対して前記復号単位時間間隔で音声パケットを供給しつづけるために挿入するものであるから、当該挿入のタイミングを検出するための下位閾値は第 1 ~ 第 4 の実施形態の下位閾値 T L のように、前記最先位置に設定するのが普通である。

【 0 1 6 7 】

ところが、下位閾値を最先位置に設定すると、補完パケット P P を挿入する位置が当該最先位置に固定されてしまうために、キュー長が 0 となるときに機械的に補完パケット P P を挿入するしかなく、キュー長が 0 の状態が 2 以上の復号単位時間にわたって継続すれば、結果的に、音声パケット P O の系列上に連続して補完パケット P P を挿入してしまふことになり、復号音声出力の品質を著しく劣化させてしまう。

【 0 1 6 8 】

これに対し本実施形態では、前記下位閾値 T L に対応する下位閾値 T L 1 を最先位置よりも上位の位置、例えば、図 8 に示すように、キュー長が 6 パケットと 7 パケットの中間の位置に設定する。

【 0 1 6 9 】

この場合、図 3 中に点線で示した音声パケット P 7 が蓄積されれば補完パケット P P の挿入は行われないが、ジッタの影響などによって当該音声パケット P 7 の受信がおくれ、蓄積がなければ、キューは音声パケット P 1 ~ P 6 だけによって構成されるようになり、キュー長は下位閾値 T L 1 を下回る。

【 0 1 7 0 】

キュー長検出部 8 2 は、前記検出信号 D 1 に対応する検出信号 D 2 を用いて下位閾値 T L 1 とキュー長との関係を監視しているから、キュー長が下位閾値 T L 1 を下回ったことを検出すると、走査読み出し部 8 3 に供給している制御信号 C 2 3 を非能動状態から能動状態に切替えるとともに、補完パケット挿入器 1 9 A に供給している制御信号 C 2 を非能動状態から能動状態に切替える。

【 0 1 7 1 】

制御信号 C 2 3 が非能動状態から能動状態に切替わったことを検出した走査読み出し部 8 3 は、その時点でキューを構成している音声パケット（図示の例では音声パケット P 1 ~ P 6 ）を走査信号 S C 3 として読み出し、当該走査信号 S C 3 を補完パケット挿入器 1 9 A に供給する。

【 0 1 7 2 】

走査信号 S C 3 を受け取った補完パケット挿入器 1 9 A は当該走査信号 S C 3 を有音・無音判定器 2 1 A に供給し、有音・無音判定器 2 1 A は、前記有音・無音判定器 2 1 と同様

10

20

30

40

50

に、走査信号 S C 3 を構成している各音声パケット P 1 ~ P 6 に対して有音・無音判定を行い、その判定結果 D C 3 を補完パケット挿入器 1 0 A に返す。

【 0 1 7 3 】

補完パケット挿入器 1 9 A は当該判定結果 D C 3 をもとに補完パケット P P を挿入する位置を決定する。なお、1 回の挿入では 1 個の補完パケット P P を挿入する。

【 0 1 7 4 】

また、挿入位置の決定にあたって補完パケット挿入器 1 0 A は、できるだけ連続しないように分散して挿入し、なおかつ、できるだけ無音区間の直後のみに補完パケット P P を挿入するように制御する。

【 0 1 7 5 】

分散して挿入するために、挿入した補完パケットには印（補完パケット・マーク）を付加し、次回に補完パケットを挿入する際には、同補完パケットの前後に挿入しないようにする。

【 0 1 7 6 】

これにより、復号音声出力の有音区間における音声の途切れを防止し、補完パケット P P の挿入にともなう音声出力の品質低下を抑制することができる。

【 0 1 7 7 】

なお、一般に音声のパケット単位で復号する装置では、一定の前記復号単位時間ごとにパケットを復号処理することが必要であるが、本実施形態により、キューを構成している音声パケットの有音・無音判定および、補完パケット P P の挿入を先行して処理することが可能となる。すなわち、処理装置の負荷が低いときに、これら进行处理することが可能であるため、処理装置の稼働率を向上させることができる。

【 0 1 7 8 】

また一般的には、補完パケット P P の挿入はそれ自体、復号音声出力の品質を劣化させる可能性があるため、通常は極力おこなわないようにすることが望ましいと考えられる。そして発明が解決しようとする課題に記載したように最先位置にのみ挿入する場合には、キュー長が 0 になったときのみ挿入するので、補完パケット P P の挿入頻度をできるだけ低減する方法であるということが出来る。

【 0 1 7 9 】

これに対し本実施形態では、キュー長が 0 ではない下位閾値 T L 1 を下回った時点で補完パケット P P を挿入するため、挿入頻度は高まる可能性が高く、場合によっては挿入する必要のない補完パケット P P を挿入してしまう可能性もある。

【 0 1 8 0 】

しかしながら本実施形態では、補完パケット P P の挿入頻度は高くても、挿入位置を選択し分散して挿入するため、挿入による復号音声出力の品質劣化はあったとしてもごくわずかであり、連続して補完パケット P P を挿入した場合などに発生し得る著しい品質劣化を防止することができる。

【 0 1 8 1 】

なお、本実施形態においても、補完パケット P P の挿入を極力おこなわないためには、前記下位閾値 T L 1 はできるだけ下位（前記最先位置に近い位置）に設定したほうがよいが、最適な挿入位置を選択し、分散して挿入するためには、できるだけ上位に設定するほうが有利である。

【 0 1 8 2 】

また、複数の復号単位時間にわたってキュー長が下位閾値 T L 1 を下回る状態が継続すると、1 復号単位時間に 1 個の補完パケット P P を挿入するとしてもキュー長が 7 パケット未満（このときキュー長は、1 復号単位時間経過するごとに 1 パケット分ずつ短縮する）のキューに対して複数の補完パケット P P を挿入しなければならなくなるため、本実施形態においても、結局は、補完パケット P P を連続的に挿入したり有音区間に挿入したりせざるを得なくなる可能性はあるが、本実施形態によれば、このような連続的挿入や有音区間に対する挿入の発生頻度を低減することが可能である。

10

20

30

40

50

【 0 1 8 3 】

なお、一般に音声 packets 単位で復号する装置では、一定の復号単位時間ごとに packets を復号処理することが必要であるが、本実施形態では、packets の有音・無音判定および、packets の挿入を先行して処理することが可能となる。すなわち、処理装置の負荷が低いときに、これらを処理することが可能であるため、処理装置の稼働率を向上させることができる。

【 0 1 8 4 】

(E - 2) 第 5 の実施形態の効果

本実施形態によれば、第 1 の実施形態の効果と同等な効果を得ることができる。

【 0 1 8 5 】

加えて、本実施形態では、補完 packets (P P) の連続的挿入や有音区間に対する挿入の発生頻度を低減することができるので、復号音声出力の品質が著しく劣化する可能性を低減することができる。

【 0 1 8 6 】

(F) 他の実施形態

なお、上記第 1 ~ 第 5 の実施形態の特徴部分は、さまざまな組み合わせで複合することができる。

【 0 1 8 7 】

例えば、第 5 の実施形態の特徴である最先位置よりも上位に設定した下位閾値 T L 1 は、第 1 の実施形態だけでなく、第 2 ~ 第 4 の実施形態と複合することが可能である。

【 0 1 8 8 】

また、第 5 の実施形態において下位閾値 T L 1 は固定されていたが、これを第 3 の実施形態の上位閾値 T H 1 や第 4 の実施形態の上位閾値 T H 2 のように、動的に変更するようにしてもよい。

【 0 1 8 9 】

なお、第 1 ~ 第 5 の実施形態では、音声 packets (P I など) には会話音声に対応するデータを収容するものとしたが、本発明は、音楽などの音声に対応するデータを収容した音声 packets を一方向に通信する場合にも適用することができる。

【 0 1 9 0 】

また、上述した上位閾値 T H は、100 packets と 99 packets の中間のキュー長に相当する位置に設定したが、現実的な実装では、30 packets 程度の位置に設定することが多いものと考えられる。ちなみに現実的なバッファメモリ 32 の容量は、200 packets 程度に設定されることが多いものと考えられる。

【 0 1 9 1 】

【 発明の効果 】

以上に説明したように、本発明によれば、バッファメモリ中に蓄積されている待ち行列を構成する複数の音声 packets が収容している音声情報の系列が持つ音声的な特性を検査した上で、当該検査結果に応じて待ち行列に対する操作を行うので、得られる音声 packets 通信の品質を向上することができる。

【 図面の簡単な説明 】

【 図 1 】 第 1 の実施形態に係る音声通信システムの主要部の構成例を示す概略図である。

【 図 2 】 第 1 の実施形態のバッファ装置の構成例を示す概略図である。

【 図 3 】 第 5 の実施形態のバッファ装置の構成例を示す概略図である。

【 図 4 】 第 4 の実施形態の音声通信装置における主要部の構成例を示す概略図である。

【 図 5 】 第 5 の実施形態に係る音声通信システムの主要部の構成例を示す概略図である。

【 図 6 】 第 1 の実施形態のバッファ装置の構成例を示す概略図である。

【 図 7 】 第 2 の実施形態に係る音声通信システムの主要部の構成例を示す概略図である。

【 図 8 】 第 3 の実施形態のバッファ装置の構成例を示す概略図である。

【 図 9 】 第 3 の実施形態に係る音声通信システムの主要部の構成例を示す概略図である。

【 符号の説明 】

10

20

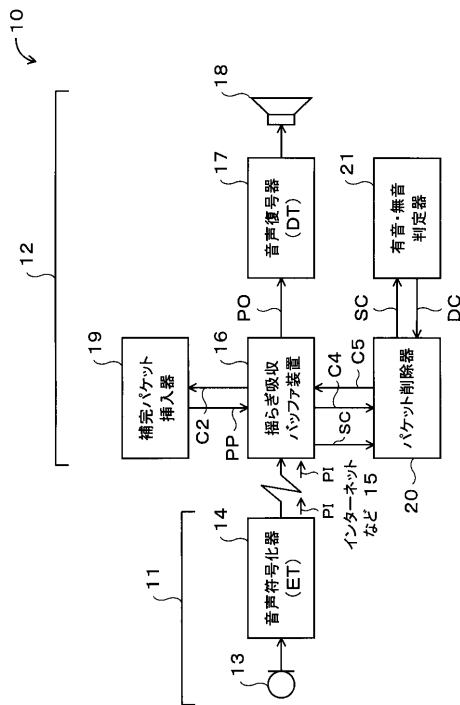
30

40

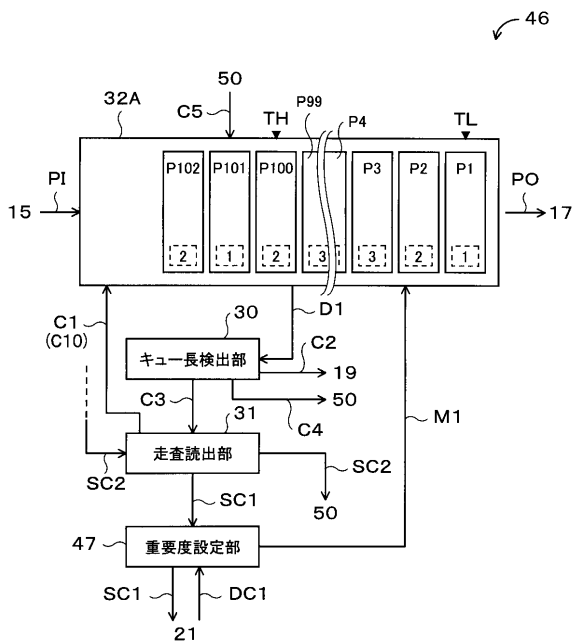
50

10, 40, 60, 80 ... 音声通信システム、11, 12, 61, 62 ... 音声通信装置、
 13, 63 ... マイクフォン、14 ... 音声復号器、15 ... インターネット、16, 46,
 66 ... バッファ装置、17 ... 音声復号器、18 ... スピーカ、19, 19A ... 補完パケット
 挿入器、20 ... パケット削除器、21, 21A ... 有音・無音判定器、30 ... キュー長検出部、
 31 ... 走査読み出し部、47 ... 重要度設定部、64 ... デュアルトーク検出部、32,
 32A ... バッファメモリ、P1 ~ P102、P0 ... 音声パケット、PP ... 補完パ
 ケット、D1、D2 ... 検出信号、C1 ~ C5 ... 制御信号、TH、TH1、TH2 ... 上位閾値
 、TL、TL1 ... 下位閾値。

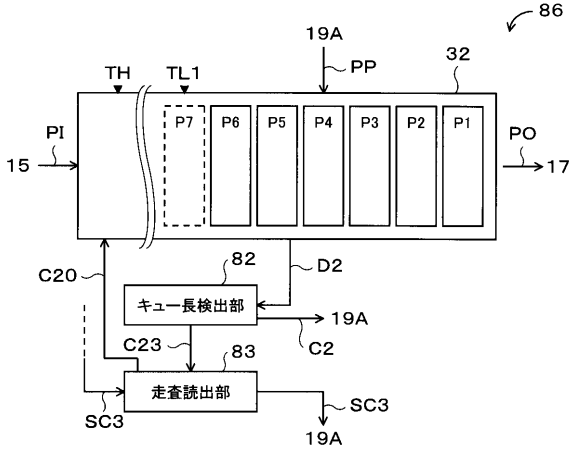
【図1】



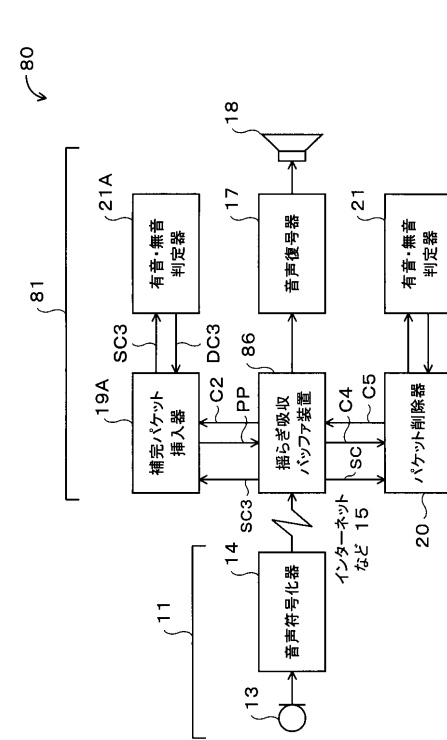
【図2】



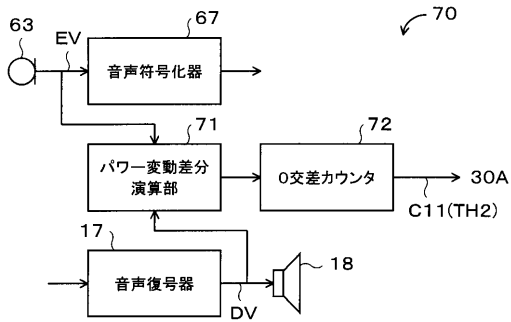
【図3】



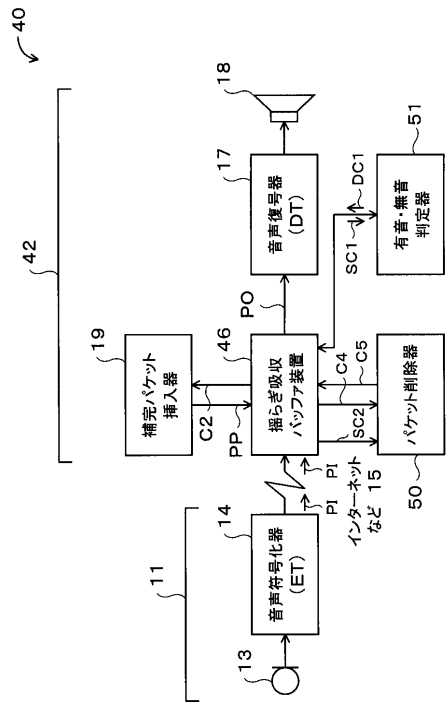
【図5】



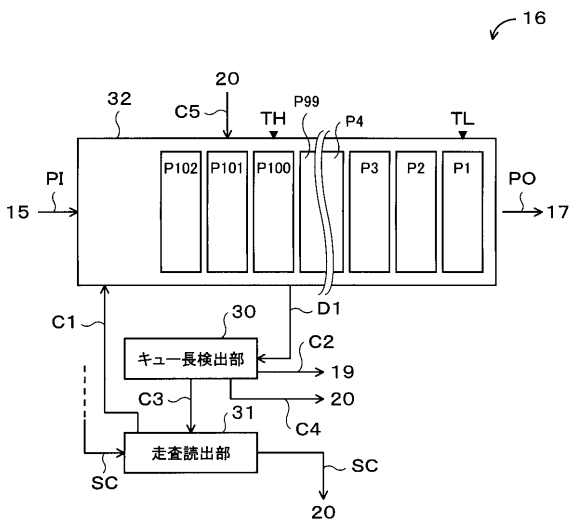
【図4】



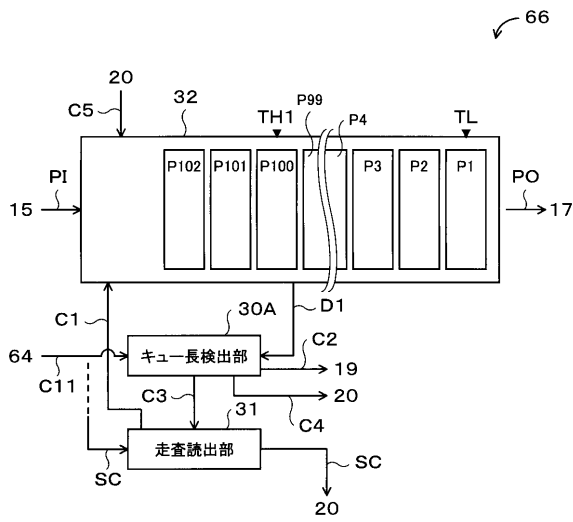
【図7】



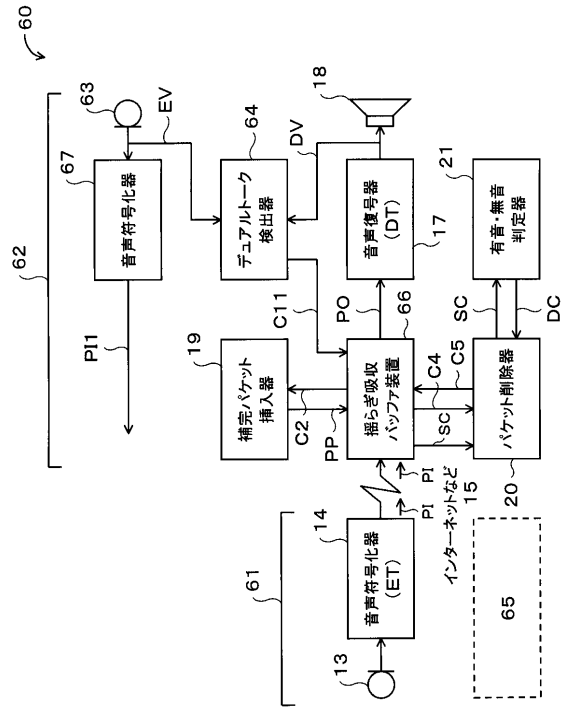
【図6】



【図8】



【図9】



フロントページの続き

(72)発明者 秋江 一良
東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

審査官 齋藤 浩兵

(56)参考文献 特開平09-083524(JP,A)
特開2000-115248(JP,A)
特開平11-163892(JP,A)
特開2000-312223(JP,A)
特開2000-099098(JP,A)
特開平02-266626(JP,A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/56