

【特許請求の範囲】**【請求項 1】**

インターネットを介して実時間音響データの送受信を行う通話装置において、
音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、
上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、
少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、
1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、
上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、
上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、
上記制御パケット生成手段は、上記管理情報と共に上記制御情報を格納した上記制御パケットを生成することを特徴とする通話装置。

10

【請求項 2】

上記制御手段は、上記受信手段が受信する上記他の通話装置からのデータパケットの伝送レートを計測し、該計測結果に基づき上記制御情報を生成することを特徴とする請求項 1 記載の通話装置。

20

【請求項 3】

上記制御手段は、ユーザ操作に基づき上記制御情報を生成することを特徴とする請求項 1 記載の通話装置。

【請求項 4】

上記制御情報は、上記他の通話装置からのデータパケットに格納される実時間音響データの圧縮率を制御する情報を有することを特徴とする請求項 1 記載の通話装置。

【請求項 5】

上記圧縮率を制御する情報は、上記他の通話装置からのデータパケットに格納される実時間音響データの圧縮すべき周波数帯域を指定する情報を有することを特徴とする請求項 4 記載の通話装置。

30

【請求項 6】

上記圧縮率を制御する情報は、上記他の通話装置からのデータパケットに格納される実時間音響データのサブバンド分割数を指定する情報を有することを特徴とする請求項 4 記載の通話装置。

【請求項 7】

上記受信手段は、受信した上記データパケットをバッファリングする受信バッファを有し、
上記制御手段は、上記受信バッファのサイズに応じて欠落データパケットの再送要求を指示する再送要求指示情報を生成し、
上記制御パケット生成手段は、上記再送要求指示情報を格納した上記制御パケットを生成することを特徴とする請求項 1 記載の通話装置。

40

【請求項 8】

上記データパケットに格納された実時間音響データは、音声とバックグラウンドミュージック及び / 又は効果音とが合成されたものであることを特徴とする請求項 1 記載の通話装置。

【請求項 9】

上記圧縮符号化手段は、上記他の通話装置からの制御パケットに格納された上記伝送レートを制御する制御情報に基づき上記他の通話装置に送信する実時間音響データの圧縮率

50

を変更する

ことを特徴とする請求項 1 記載の通話装置。

【請求項 10】

インターネットを介して実時間音響データの送受信を行う通話方法において、
音声を含む実時間音響データを圧縮符号化する圧縮符号化工程と、

上記圧縮符号化工程にて圧縮符号化されたデータを格納したデータパケットを生成する
データパケット生成工程と、

少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを
生成する制御パケット生成工程と、

1 以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する
送信工程と、

上記他の通話装置からのデータパケット及び制御パケットを受信する受信工程と、

上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制
御工程とを有し、

上記制御パケット生成工程では、上記管理情報と共に上記制御情報を格納した上記制御
パケットを生成する

ことを特徴とする通話方法。

【請求項 11】

上記制御工程では、上記受信工程にて受信する上記他の通話装置からのデータパケット
の伝送レートを計測し、該計測結果に基づき上記制御情報を生成する

ことを特徴とする請求項 10 記載の通話方法。

【請求項 12】

上記制御工程では、ユーザ操作に基づき上記制御情報を生成する

ことを特徴とする請求項 10 記載の通話方法。

【請求項 13】

上記制御情報は、上記他の通話装置からのデータパケットに格納される実時間音響デー
タの圧縮率を制御する情報を有する

ことを特徴とする請求項 10 記載の通話方法。

【請求項 14】

上記圧縮率を制御する情報は、上記他の通話装置からのデータパケットに格納される実
時間音響データの圧縮すべき周波数帯域を指定する情報を有する

ことを特徴とする請求項 13 記載の通話方法。

【請求項 15】

上記圧縮率を制御する情報は、上記他の通話装置からのデータパケットに格納される実
時間音響データのサブバンド分割数を指定する情報を有する

ことを特徴とする請求項 13 記載の通話方法。

【請求項 16】

上記制御工程では、受信した上記データパケットをバッファリングする受信バッファの
容量に応じて欠落データパケットの再送要求を指示する再送要求指示情報を生成し、

上記制御パケット生成工程では、上記再送要求指示情報を格納した上記制御パケットを
生成する

ことを特徴とする請求項 10 記載の通話方法。

【請求項 17】

上記データパケットに格納された実時間音響データは、音声とバックグラウンドミュー
ジック及び / 又は効果音とが合成されたものである

ことを特徴とする請求項 10 記載の通話方法。

【請求項 18】

上記圧縮符号化工程では、上記他の通話装置からの制御パケットに格納された上記伝送
レートを制御する制御情報に基づき上記他の通話装置に送信する実時間音響データの圧縮
率を変更する

ことを特徴とする請求項 10 記載の通話方法。

【請求項 19】

2 以上の通話装置がインターネットを介して実時間音響データの送受信を行う通話システムにおいて、

各通話装置は、

音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、

上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、

少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、

1 以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、

上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、

上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、

少なくとも 2 つの通話装置の一方の上記制御パケット生成手段は、上記管理情報と共に他方の通話装置からのデータパケットの伝送レートを制御する上記制御情報を格納した上記制御パケットを生成し、

上記他方の通話装置の上記圧縮符号化手段は、上記一方の通話装置からの制御パケットに格納された上記制御情報に基づき該一方の通話装置に送信する実時間音響データの圧縮率を変更する

ことを特徴とする通話システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、例えば V o I P (Voice over Internet Protocol) を使用したいわゆるインターネット電話等を行う通話装置、通話方法及び通話システムに関し、特に高音質の音声及び音響からなる実時間音響データをやり取りするために好適な通話装置、通話方法及び通話システムに関する。

【背景技術】

【0002】

音声を IP パケットにしてカプセル化することで IP 網を介した音声通話を可能とする技術として、V o I P がある。V o I P による通話を行うためには、通話したい相手の情報の取得、通話したい相手の呼び出し、応答といった一連の情報交換をする必要があり、これらの目的のために、S I P (Session Initiation Protocol) 等の呼制御プロトコルが使用される。

【0003】

近年、このような V o I P を含め、インターネット等の通信回線網を利用し、実時間動画像データや、音声データ等の実時間音響データ等の実時間データを送信するシステムが増えている。このようなインターネット等の公衆回線網においては、複数の利用者がネットワークの帯域を供給しているため、輻輳制御手法、すなわち輻輳の回避及び輻輳発生時の鎮静化手法が大きな課題となっている。従って、実時間性を重要とする通信形態が増えるという変化に伴い、実時間データの通信における輻輳制御手法が重要になってきている(下記特許文献 1 参照)。

【0004】

ところで、従来、V o I P を使用した実時間音声コミュニケーションにおいては、人の音声の主たる情報が含まれているのは、4 ~ 5 KHz 付近の周波数帯域であったため、チャンネル数は 1、標本化周波数を 8 KHz ~ 16 KHz に設定するのが常道であった。しかし、このような狭い帯域では背景雑音や雰囲気等を伝達するのは困難である。

10

20

30

40

50

【0005】

そこで、チャンネル数を2以上に設定し、標本化周波数をCDDA (Compact Disk Digital Audio) 並の44.1 KHzまで引き上げると、背景雑音や雰囲気等も伝送でき、また加えて、音声コミュニケーションに、既存の楽曲コンテンツを流用した高音質なBGM機能等を付加することなども可能になる。

【0006】

ところがチャンネル数を2以上に設定し、標本化周波数をCDDA並の44.1 KHzまで引き上げて、無圧縮の16 bits Linear PCM (Phase Code Modulation: パルス符号変調) のデータを伝送する場合、ネットワーク上の伝送ビットレートは1411.2 Kbps以上になってしまい、実インターネット環境で利用するのは難しい。したがって、現状では高能率符号化方式を用いる必要があるが、近時のCPUの性能の向上により、実時間音声コミュニケーションにおいても高能率符号化方式を採用することが可能になってきている。

10

【0007】

なお、高能率符号化方式にはいくつかの方式があるが、まず周波数分解を行い、信号を複数のサブバンドに分割したうえで、それぞれのブロックを、聴覚心理特性を利用して符号化精度を適応的に変化させて、必要最低限のビット数で所定の周波数成分をエントロピー符号化するものが殆どである。この高能率符号化により、符号化する周波数成分に応じてビットレートは可変になる。

【0008】

即ち、従来の一一般の電話のように、伝送ビットレートも音のクオリティも低い場合には、ネットワーク状況に応じて動的に圧縮方式のパラメータを変更して輻輳制御をする必要が生じなかったものの、BGM機能等を付加したような高音質な実時間音響データをやり取りする実時間音声コミュニケーションを行う場合には、ネットワークのデータ伝送レートが高くなる。したがって、通話中に何らかの要因でネットワーク帯域が混雑しはじめた場合、通話が一時的に遮断されることを余儀なくされることが少なくなく、このようなBGM機能が付加されているような高音質音声の実時間コミュニケーションにおいては、上記特許文献1のように、動的に圧縮方式のパラメータを変更する必要が生じる。

20

【0009】

【特許文献1】特開2001-320044号公報

30

【発明の開示】

【発明が解決しようとする課題】

【0010】

しかしながら、VoIPによる実時間音声コミュニケーションにおいては、基本的にはSIP等の呼制御プロトコルを用いてシグナリングを行い、通話相手のIPアドレス・ポート番号・音声及び音響圧縮方式を知る。そして、自身とその目標アドレスとの間で双方向に RTP (Real-Time Transport Protocol) 及び RTCP (Real-Time Control Protocol) のストリームを送受信するのが基本的な仕組みである。通常、一度 RTP のストリームが確立されると、音声圧縮方式は再呼制御を行わない限り変更できない。また、再呼制御を行うと通話が一時的に遮断される。

40

【0011】

即ち、符号化する周波数成分に応じて別々の音声・音楽圧縮方式として分類してしまうと、再呼制御が必要となるため、ネットワーク帯域が何らかの理由で混雑する等した際、そのままの伝送ビットレートであっても、伝送ビットレートを下げるために必要となる再呼制御によっても、通話が一時的に遮断されてしまうという問題点がある。そこで、同一の音声・音楽圧縮方式がネットワーク帯域に応じて、通話を遮断することなく動的にサブバンド分割数及び符号化すべき周波数成分を決定する仕組みの一般化が望まれる。

【0012】

本発明は、このような従来の実情に鑑みて提案されたものであり、音声にBGM及び効果音等を合成した高音質な音声コミュニケーションをインターネットを介して行う際に、

50

通話を遮断することなく、ネットワーク帯域に応じて符号化・伝送する音響周波数成分を動的に決定することができる通話装置、通話方法及び通話システムを提供することを目的とする。

【課題を解決するための手段】

【0013】

上述した目的を達成するために、本発明に係る通話装置は、インターネットを介して実時間音響データの送受信を行う通話装置において、音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、上記制御パケット生成手段は、上記管理情報と共に上記制御情報を格納した上記制御パケットを生成することを特徴とする。

10

【0014】

本発明においては、他の通話装置からのデータパケットの伝送レートを制御する制御情報生成し、これを上記管理情報と共に格納した上記制御パケットとして他の通話装置に送信することで、これを受信する他の通話装置に対し、制御パケットの制御情報に基づき他の通話装置が送信するデータパケットの伝送レートを設定させることができる。即ち、ネットワーク帯域が急激に混雑し始めた場合や受信側で聞き取りにくい等の状態が生じた場合に、データパケットの受信側でデータパケットの送信側の送信データの伝送レートの制御をすることができ、またこの制御情報を、呼制御の際のメッセージではなく、制御パケットと共に送信するため、データパケットの送受信を停止する必要がない。

20

【0015】

また、上記制御手段は、上記受信手段が受信する上記他の通話装置からのデータパケットの伝送レートを計測し、該計測結果に基づき上記制御情報を生成するか、又はユーザ操作に基づき上記制御情報を生成することができ、この制御情報としては、上記他の通話装置からのデータパケットに格納される実時間音響データの圧縮率を制御する情報、例えば上記他の通話装置からのデータパケットに格納される実時間音響データの圧縮すべき周波数帯域及び/又はサブバンド分割数を指定する情報等とすることができ、これにより、他の通話装置が送信してくるデータパケットに格納する実時間音響データの圧縮符号化方式を変更させることができる。

30

【0016】

更に、上記受信手段は、受信した上記データパケットをバッファリングする受信バッファを有し、上記制御手段は、上記受信バッファのサイズに応じて欠落データパケットの再送要求を指示する再送要求指示情報を生成し、上記制御パケット生成手段は、上記再送要求指示情報を格納した上記制御パケットを生成することができ、他の通話装置からの伝送レートを制御する制御情報と共に、受信されるまでにエラーが生じるか又は伝送途中で失われる等した欠落パケットの再送要求を行う再送要求指示情報を制御パケットにて送信することができる。

40

【0017】

更にまた、上記データパケットに格納された実時間音響データは、音声とバックグラウンドミュージック及び/又は効果音とが合成されたものであり、音声のみではなくバックグラウンドミュージック等が付加されて情報量が多い送信データであっても必要に応じて伝送レートを制御することにより通信が遮断されることがない。

【0018】

また、上記圧縮符号化手段は、上記他の通話装置からの制御パケットに格納された上記伝送レートを制御する制御情報に基づき上記他の通話装置に送信する実時間音響データの圧縮率を変更することができ、通話相手となる他の通話装置からの要求に応じて送信する

50

データパケットに格納する実時間音響データの圧縮率を変更し、送信するデータパケットの伝送レートを可変とすることができる。

【0019】

本発明に係る通話方法は、インターネットを介して実時間音響データの送受信を行う通話方法において、音声を含む実時間音響データを圧縮符号化する圧縮符号化工程と、上記圧縮符号化工程にて圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成工程と、少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成工程と、1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信工程と、上記他の通話装置からのデータパケット及び制御パケットを受信する受信工程と、上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御工程とを有し、上記制御パケット生成工程では、上記管理情報と共に上記制御情報を格納した上記制御パケットを生成することを特徴とする。

10

【0020】

本発明に係る通話システムは、2以上の通話装置がインターネットを介して実時間音響データの送受信を行う通話システムにおいて、各通話装置は、音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、少なくとも2つの通話装置の一方の上記制御パケット生成手段は、上記管理情報と共に他方の通話装置からのデータパケットの伝送レートを制御する上記制御情報を格納した上記制御パケットを生成し、上記他方の通話装置の上記圧縮符号化手段は、上記一方の通話装置からの制御パケットに格納された上記制御情報に基づき該一方の通話装置に送信する実時間音響データの圧縮率を変更することを特徴とする。

20

【0021】

本発明においては、一方の通話装置が、自身の通話対象となる他方の通話装置が送信するデータパケットの伝送レートを制御することができ、また、他方の通話装置からの要求に応じて該他方の通話装置に送信する実時間音響データの圧縮符号化方式を変更して当該他方の通話装置に伝送するデータパケットの伝送レートを制御することができる。

30

【発明の効果】

【0022】

本発明に係る通話装置によれば、インターネットを介して実時間音響データの送受信を行う通話装置において、音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、上記制御パケット生成手段は、上記管理情報と共に上記制御情報を格納した上記制御パケットを生成するので、データパケットの受信側で、制御パケットにより、データパケットの送信側となっている他の通話装置の送信伝送レートの制御をすることができ、例えばネットワーク帯域が急激に混雑し始めた場合や受信側で聞き取りにくい等の状態が生じた場合には通話相手となる他の通話装置が送信する伝送レートを再呼制御することなく下げさせることができ、よって高音質で情報量が大きい実時間音響データを送受信するような場合であっても通話を継続することができる。

40

【0023】

50

また、本発明に係る通話システムによれば、2以上の通話装置がインターネットを介して実時間音響データの送受信を行う通話システムにおいて、各通話装置は、音声を含む実時間音響データを圧縮符号化する圧縮符号化手段と、上記圧縮符号化手段により圧縮符号化されたデータを格納したデータパケットを生成するデータパケット生成手段と、少なくとも上記データパケットの送受信を管理する管理情報を格納した制御パケットを生成する制御パケット生成手段と、1以上の他の通話装置に対して、上記データパケット及び上記制御パケットを送信する送信手段と、上記他の通話装置からのデータパケット及び制御パケットを受信する受信手段と、上記他の通話装置からのデータパケットの伝送レートを制御する制御情報を生成する制御手段とを有し、少なくとも2つの通話装置の一方の上記制御パケット生成手段は、上記管理情報と共に他方の通話装置からのデータパケットの伝送レートを制御する上記制御情報を格納した上記制御パケットを生成し、上記他方の通話装置の上記圧縮符号化手段は、上記一方の通話装置からの制御パケットに格納された上記制御情報に基づき該一方の通話装置に送信する実時間音響データの圧縮率を変更するので、一方の通話装置が、自身の通話対象となる他方の通話装置からのデータパケットの伝送レートを制御することができ、また、他方の通話装置からの要求に応じて他方の通話装置に伝送するデータパケットに格納する実時間音響データの圧縮符号化方式を変更して伝送レートを制御することができ、情報量が多い、例えばBGMや効果音等が付加された等高音質な実時間音響データをやり取りするような場合であっても伝送レートを動的に変化させることができるので、ネットワーク帯域が混雑した場合でも通話が遮断されることがない。

10

20

【発明を実施するための最良の形態】

【0024】

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、2以上の通話装置がインターネットを介してVoIPにより通話を行う通話装置としてのVoIPクライアント及びこれを備えた通話システムとしてのVoIPシステムに適用したものである。本実施の形態におけるVoIPクライアントは、通話装置間での通話継続を可能とするため、ネットワーク帯域に応じて、音質を可変にする、即ち同一の音声・音楽圧縮方式におけるサブバンド分割数及び符号化すべき周波数成分を動的に決定するものである。

【0025】

先ず、本実施の形態におけるVoIPを使用したネットワークコミュニケーションを行うVoIPシステムの概略について説明する。図1は、本実施の形態におけるVoIPシステムの一例を示す模式図である。本実施の形態におけるVoIPシステムでは、例えば2チャンネル以上で、且つユーザ間で、通話のみではなく様々な効果音及びBGM等も共有することができる高音質の音声コミュニケーションを実現するものである。なお、本実施の形態におけるVoIPシステムは、2つの通話装置（以下、VoIPクライアントという。）間で行なわれるものとするが、VoIPシステムを構成するVoIPクライアントは2つに限らず、従ってVoIPクライアントを介してネットワークコミュニケーションに参加可能な参加者は2以上であってもよい。

30

【0026】

図1に示すように、VoIPシステム100は、例えばPC(Personal Computer)等のVoIPクライアント111と、これとインターネット130を介して接続されたVoIPクライアント121とを有する。

40

【0027】

このVoIPシステム100においては、VoIPクライアント111のユーザ110と、VoIPクライアント121のユーザ120とは、自身のVoIPクライアント111、121に搭載される後述するVoIP用のアプリケーション(ソフト・フォン)等と、例えばマイクロフォンとヘッドフォンとからなるヘッドセット又はマイクロフォンと受話器とからなるハンドセットとを使用し、インターネット130を介して通信相手とコミュニケーションを行う。

50

【0028】

インターネット130は、一般公衆回線等の通信回線や、情報通信ネットワークを複数接続することによって世界中に広がったネットワーク環境である。現在、広帯域、高速な通信回線の普及によってブロードバンド伝送(Broadband Transmission)を可能としている。光ファイバー、非対称デジタル加入者線、無線等を用い、500kbps以上の通信回線でネットワークを構成している。

【0029】

このインターネット130には、VoIP通信を制御するVoIPサーバ131、及び音源データ132及びダウンロードユーザインフォメーション133等のデータを管理するウェブサーバ134等が接続されている。また、各VoIPクライアント111、121には、各VoIPクライアント111、121が有するウェブブラウザ112、122等によりウェブサーバ134からダウンロードするか、又は自身で購入若しくは編集した音源データ113、123が記憶されている。

【0030】

ウェブサーバ134のデータベースに格納されている音源データ132、及びユーザがダウンロード等して所持している音源データ113、123は、例えばBGM(Back Ground Music)等に使用される音楽や、波の音・拍手の音・雷鳴・ベルの音等の各種効果音であり、これらの音源データは、ユーザ間のネットワークコミュニケーションにて使用することができる。

【0031】

次に、このVoIPクライアントについて説明する。図2は、VoIPシステムを構成するVoIPクライアントの機能を示すブロック図である。図2に示すように、このようなインターネットコミュニケーションを行うためのVoIPクライアント20は、コミュニケーションに参加している参加者、即ち通話相手へデータを送信する送信手段21と、通話相手からのデータを受け取る受信手段41とを有する。

【0032】

送信手段21は、マイクが接続され、外部(ユーザ)の音声をキャッチするマイクキャプチャ(MIC capture)22と、例えばMP3(MPEG(Moving Picture Experts Group)1オーディオLayer3)、又はMPEG4等に圧縮された各種効果音(Sound Effect: SE)の音源ファイルが記憶された効果音ファイル記憶部23と、同じく圧縮された各種BGMの音源ファイルが記憶されたBGMファイル記憶部24と、効果音ファイル及びBGMファイルを読み出しデコードする夫々デコーダ25及び26と、マイクによりキャプチャした音声、効果音、及びBGMのゲインを制御して音量調整するゲイン調整部27~29と、これらの3つの音を合成する合成部30と、合成した音を圧縮符号化するエンコーダ31と、圧縮符号化されたデータをデータパケットとしてのRTPパケットに格納するデータパケット生成手段としてのパケット化部(packetize)32aと、後述する制御情報を生成する制御部34と、該制御情報及びRTPパケットをコントロールする管理情報等をRTPパケットに格納する制御パケット生成手段としてのRTPパケット化部32bと、RTPパケット及びRTPパケット等を送信する送信部33とを有する。送信部33から送られたRTPパケットがインターネット130を介して通信対象となる他のVoIPクライアントに送信される。

【0033】

一方、インターネット130を介して通信相手の送信手段21から送られるデータを受信する受信手段41は、RTPパケット及びRTPパケット等を受信する受信部42と、受信したRTPパケットをデパケッタイズするRTPデパケット化部(depaketize)43aと、RTPパケットをデパケッタイズするRTPデパケット化部(depaketize)43bと、RTPパケットの到着時間を補正するデジッタ部(de-jitter)44と、送られたRTPパケットのエラーが生じた部分又は伝送中に損失した部分等の欠落部分を補償するパケット補償部(packet loss compensator)45と、パケット補償部45からのデータを復号伸張するデコーダ46と、例えばMP3、MPEG4等に圧縮された着信音

10

20

30

40

50

の音源ファイルが記憶された着信音ファイル (Ring Tone File) 記憶部 47 と、着信音ファイルを読み出しデコードするデコーダ 48 と、デコーダ 46 及び 48 のゲイン調整をする夫々ゲイン調整部 49 及び 50 と、上述した送信手段 21 における合成部 30 にて合成された送信用の PCM データ、即ち送信者側自身の音源のゲイン調整するゲイン調整部 52 と、ゲイン調整された送信用のデータ、通信相手から送信されてきたデータ、及び着信音データを合成する合成部 53 と、合成部 53 にて合成されたデータをヘッドホン (HP) へ出力する出力部 54 と、合成部 53 へ出力される着信音とは別に着信音のゲインを調整するゲイン調整部 51 と、ゲイン調整された着信音を外部へ出力するスピーカ (SP) 55 とを有する。

【0034】

10

次に、この VoIP クライアントのデータの送受信方法について説明する。まず、送信側において、マイクキャプチャ 22 は、マイクにより入力されたユーザの音声をキャッチしてゲイン調整部 27 へ送る。ゲイン調整部 27 は、音声データに対し、ユーザの指示によるか又は自動的に、ゲイン係数 k_1 を乗算し、所望の大きさにゲイン調整する。

【0035】

効果音ファイル記憶部 23 及び BGM ファイル記憶部 24 は、例えば MP3、MP EG 4 等の圧縮技術により予め圧縮された音源ファイルが記憶された、例えばハードディスクドライブ (HDD)、ROM (read only memory) 又は光磁気ディスクからなり、デコーダ 25 及び 26 は、これらの圧縮符号化データを読み出し、読み出した圧縮符号化データを PCM データに変換する。更に、デコーダ 25、26 は、変換した PCM データを夫々ゲイン調整部 28、29 へ送り、ゲイン調整部 28、29 は、送られてきたデータに対し、ユーザの指示によるか又は自動的に、夫々ゲイン係数 k_2 及び k_3 を乗算して、所望の大きさにゲイン調整する。

20

【0036】

ゲイン調整部 27 ~ 29 にてゲイン調整された実時間音響データの音声、BGM 及び効果音は、合成部 30 に供給され、合成部 30 はゲイン調整部 27 ~ 29 の出力を飽和処理しつつ加算し、この加算結果をエンコーダ 31 に出力する。エンコーダ 31 は、この加算結果をネットワークにて送信するために、例えば MP EG 4 等にエンコードする。エンコードされたデータは、リアルタイム・トランスポート・プロトコル (Real-time Transport Protocol: RTP) に従ってデータをパケット化する RTP パケット化 (packetize) 部 32 に供給される。

30

【0037】

RTP パケット化部 32 a は、エンコードデータを RTP パケットにパケット化し、パケット化データは、送信部 33 からインターネット 130 を介して通信相手の VoIP クライアントへ送信される。

【0038】

また、制御部 34 は、後述する RTP パケットのレポートブロック等に格納される各種管理情報を管理すると共に、RTP パケットの拡張部分に格納される制御情報を生成する。制御情報とは、通信相手となる他の VoIP クライアントが送信してくる RTP パケットに格納される実時間音響データの圧縮率を指定する例えばエンコード帯域及びサブバンド分割数等の情報、及び伝送途中でエラーが生じたり又は失われたような欠落パケットの再送要求を指示する情報等である。また、本実施の形態における VoIP クライアントが送受信する RTP パケットの拡張部分には、RTP 受信部 42 にて受信したパケットのジッタを相殺するための揺らぎ吸収バッファ及び再生する前の無圧縮の PCM データを格納しておく PCM バッファの容量等も記述することができる。

40

【0039】

そして、RTP パケット化部 32 b は、この制御情報及び管理情報を RTP パケットにパケット化し、このパケット化データも送信部 33 からインターネット 130 を介して通信相手の VoIP クライアントへ送信される。

【0040】

50

また、V o I Pによる通話では、R T Pパケットを送信する前に、制御部34がS I P等の呼制御プロトコルにより、送信部33を介して、通信対象となる他のV o I Pクライアントに対し呼シグナリングを行う。そして、他のV o I Pクライアントの間でR T P / R T C Pセッションが確立された後、送信部33がR T Pパケット及びR T C Pパケットを送信することができる。呼シグナリング方法及びR T P / R T C Pパケットの詳細は後述する。

【0041】

一方、受信側においては、受信部42がインターネット130を介して送られてくるR T Pパケット及びR T C Pパケット等を受信する。そして、R T C Pデパケット化部43bは、受信したR T C Pパケットを分解し、制御部34は、R T C Pパケットの拡張部分に格納されている制御情報を取り出し、この制御情報のうち、上述したエンコード帯域及びサブバンド分割数等からなる圧縮率制御情報をエンコーダ31に供給する。エンコーダ31は、通話相手となる他のV o I Pクライアントからの要求としての上記圧縮制御情報により、エンコードする実時間音響データの圧縮率を変更する。

10

【0042】

また、デパケット化部43は、受信部42が受信したR T Pパケットを逆パケット化した後、デジッタ部44が例えばネットワークの状態等により受信が遅れる等して到着時刻が不均等になっているデータを補正し、等間隔化する。この補正は、逆パケット化されI P、U D P (User Datagram Protocol) から分解されたR T Pのタイムスタンプ、シーケンシャルナンバーを基に行なわれる。その後、パケット補償部45がネットワークの送受信において欠落又は受信不能等となったパケットを保障する処理、具体的には、欠落したパケットの代わりにその前又は後のパケットと同じパケットを使用するようにしたり、欠落したパケットの再送要求をして改めて欠落パケットを受信するようにする等したりしてパケットの損失を補償する。こうして得られたM P E G 4等の圧縮データはデコーダ46にてデコードされ、これをゲイン調整部49がユーザ指示又は自動的に、ゲイン係数k 5を乗算してゲイン調整する。

20

【0043】

デコーダ48は、他のV o I Pクライアントから呼シグナリングされたとき、即ち電話がかかってきたときにユーザに通知するための呼び出し用の呼び出し音又は呼び出し用使用する音楽等からなる着信音の音源ファイルを着信音ファイル記憶部47から読み出す。ファイル記憶部47からの着信音データは、ユーザの所望によって予め選択されており、着信のタイミングに従って図示しないR A Mに読み出されながらデコーダ48にてデコードされる。着信音の音源ファイルも、B G M等の音源ファイル等と同様にM P 3、M P E G 4等に圧縮されたものとなっており、ファイル単位の着信音データとしてファイル記憶部47に複数ファイル分記憶されている。そして、デコーダ48は、読み出した音源ファイルをデコードし、デコードされた音源データをゲイン調整部50、51がユーザ指示又は自動的に、夫々ゲイン係数k 6、k 7を乗算することでゲイン調整する。

30

【0044】

合成部53は、着信音及び通信相手から送信されゲイン調整部49にてゲイン調整された、通話相手から受け取った受信側実時間音響データと、ゲイン調整部52から出力される自身が送信する送信側実時間音響データとを加算処理する。ゲイン調整部52は、送信される送信側実時間音響データを通話相手と共有するため、送信側実時間音響データにユーザが設定するループバック音量レベルであるゲイン係数k 4を乗算するものである。

40

【0045】

そして、合成部53にて合成されたこれらのデータは、ヘッドフォン54を介して出力されユーザに伝えられる。また、着信音ファイル記憶部47から読み出された着信音データは、ヘッドフォン54とは別にスピーカ55からも出力されるよう構成されている。

【0046】

このようなV o I Pクライアントにおいては、R T C Pパケットに制御情報を格納して通話相手となる他のV o I Pクライアントに送信することで、ネットワークの輻輳状態に

50

応じて、受信する RTP パケットの伝送レートを制御することができる。即ち、他の V o I P クライアントが圧縮すべきエンコード帯域及びサブバンド分割数等を指定し、例え一時的に音質が落ちたとしても、通話を継続させるようにすることができる。

【 0 0 4 7 】

更に、一般の電話に比してダイナミックレンジを広くすることで、ステレオ等の高音質の音声と B G M 及び効果音等からなる音響とが合成された合成音を送受信することができ、従って B G M 等を付加しても会話（音声）をマスキングすることがない。また、効果音及び B G M 等の音源ファイルから読み出した音と、通信者自身が話した音声、即ちマイク音とを別々にゲイン調整することができるため、効果音及び B G M の音量を所望のレベルに調整することができ、ユーザは、通信相手へ例えば自身の気分を伝えたりすることができる。

10

【 0 0 4 8 】

また、着信音をヘッドホン 5 4 とスピーカ 5 5 とで別々に出力することにより、例えばユーザがヘッドホン 5 4 を使用して V o I P による通信中に一時的に V o I P クライアント 2 0 から離れたり、通話が終了した後もヘッドホン 5 4 をセットしたままにしておいたりした場合であっても、着信音がスピーカ 5 5 から外部に出力されるようにすることができる。

【 0 0 4 9 】

次に、この V o I P クライアントに使用されるソフトウェアについて説明する。図 3 は、一般的な P C 向け V o I P クライアントアプリケーションを含むソフトウェアモジュールを示す図である。V o I P クライアントは、この図 3 に示す開放型システム間相互接続（Open System Interconnection：O S I）のアーキテクチャに基づく各階層のプロトコルに応じたソフトウェアモジュール 1 を実行することにより上述の図 2 に示した機能を実現する。

20

【 0 0 5 0 】

図 3 において下位層から上位層に向かって各階層を説明する。まず、物理層 2 としての機能にはユニバーサル・シリアル・バス（Universal Serial Bus：U S B）カメラドライバ 2 a、U S B オーディオドライバ 2 b 及び各種ドライバ 2 c がある。カメラドライバ 2 a からのビデオデータやオーディオドライバ 2 b からのオーディオデータの伝送条件の物理的条件を合わせるレイヤである。

30

【 0 0 5 1 】

次に、データリンク層としての機能には、オペレーティングシステム（Operating System：O S）3 がある。隣接ノード間の誤りのないデータ伝送を実行するためのものである。

【 0 0 5 2 】

そして、ネットワーク層としての機能には、インターネットプロトコル（Internet Protocol：I P）制御部 4 がある。ネットワーク層は、データ送受信に使用する通信経路を選択し、フロー制御・品質制御などの通信制御を行うところである。信頼性を追求しないコネクションレス（Connectionless）パケット伝送プロトコルである I P は、信頼性保証機能、フロー制御機能、エラー回復機能を上位階層（トランスポート層とアプリケーション層）に任せている。

40

【 0 0 5 3 】

トランスポート層としての機能には、トランスポート・コントロール・プロトコル（Transport Control Protocol：T C P）/ ユーザ・データグラム・プロトコル（User Datagram Protocol：U D P）制御部 5 がある。

【 0 0 5 4 】

トランスポート層では、I P アドレスを使用してエンド・ツー・エンドの伝送を行う。ネットワークの種類に依存せず、要求される品質クラスに従ってフロー制御や順序制御を行う。T C P は信頼性保証機能を持ち、伝送したデータの各バイトにシーケンス番号を付け、受信側から受け取り通知（Acknowledgment：A C K（確認応答））が送られてこなけ

50

ればデータを再送する。UDPは、アプリケーション間のデータグラムの送信機能を提供する。IPネットワークを用いて、音声・動画像をストリーミング再生する場合、一般にエラー時に再送を行うTCPのようなトランスポート・プロトコルは使用できない。また、TCPは、1対1通信用のプロトコルであり、複数の相手に情報を送信することができない。そこで、このような用途には、UDPが用いられる。即ち、例えば再送制御を必要とする等、信頼性が高い通信を行う際にはTCP通信としてTCPパケットを生成し、音声及び映像データ、並びに後述するSIPを伝送する等、リアルタイム性が高い通信を行う際にはUDP通信としUDPパケットを生成する。

【0055】

UDPは、アプリケーションのプロセスがリモートマシン上の他のアプリケーションのプロセスへデータを伝送することを、最小のオーバーヘッドで行えるように設計されている。そのため、UDPのヘッダに入る情報は、送信元ポート番号、宛先ポート番号、データ長、チェックサムのみであり、TCPにあるパケットの順序を表す番号を入れるフィールドがないので、ネットワーク上で異なる経路を介して伝送されるなどによりパケットの順序が入れ替わってしまった場合に、その順序を正しい状態に戻す処理を行うことができない。また、送信時のタイムスタンプ等の時間情報を入れるフィールドは、TCPにもUDPにもない。

【0056】

セッション層6としての機能には、セッション・イニシエーション・プロトコル (Session Initiation Protocol: SIP) 制御部11、RTP/RTCP制御部12、コーデック (codec) 13、通話音とBGM又は効果音の合成処理ソフトウェアを構成する保留音制御部14、BGM合成部15、及び着信音制御部16がある。セッション層6は、情報の伝送制御を行う。アプリケーション間における対話モードを管理して会話単位の制御を行う。

【0057】

SIP制御部11は、IPネットワーク上でマルチメディアセッションを確立・変更・終了するための、セッション層6のシグナリングプロトコル (RFC3261で標準化) により呼制御を行う。

【0058】

また、RTP/RTCP制御部12のうち、RTPは、音声データにBGM及び効果音等が合成され圧縮符号化された送信データを送信するためのプロトコルであり、送信データに時間情報及び順序情報を付与してネットワークを通じて音声データ送受信する機能を有する。また、RTCPは、RTPを制御する制御プロトコルであり、RTPのフロー制御、クロック同期及びデータの再生時刻の認識、情報源の認識等を行う機能を有する。

【0059】

コーデック13は、送信音 (伝送データ) を後述する高能率音響圧縮及び復号する機能を有する。

また、プレゼンテーション層としての機能には、VoIP通話制御7がある。プレゼンテーション層では、アプリケーションで送受信する情報の表現形式を管理して、データの変換や暗号化を行う。VoIP通話制御部7は、VoIP通話機能の全体をコントロールする制御部である。

【0060】

最上層のアプリケーション層としての機能には、コンピュータを視覚的に操作するためのグラフィカルユーザインターフェース (Graphical User Interface: GUI) 8がある。アプリケーション層は、ユーザプログラムで使用する通信機能の外部仕様を管理して、それに基づく情報のやり取りを行う層であり、GUI8は、ユーザ操作のインターフェイスを提供し、ユーザの手入力情報をハンドリングする。

【0061】

図4は、GUIの一例を示す模式図である。GUI8は、図4に示すように、VoIPクライアントアプリケーション1の終了処理を行うアプリケーション制御部61と、情報

表示部 6 2 と、ダイヤル部 6 3 と、ヘッドセットボリューム部 6 4 a と、スピーカボリューム部 6 4 b と、サウンドエフェクト選択表示部 6 5 と、サウンドエフェクト制御部 6 6 と、BGM 選択表示部 6 7 と、BGM 制御部 6 8、音質制御部 6 9 とを有する。

【0062】

情報表示部 6 2 は、ユーザが通信対象となる相手先の番号をダイヤルした場合のダイヤル番号や、通信相手が話し中か否か等の相手状態等を表示する。ダイヤル部 6 3 は、V o I P 相手先をダイヤルするため等のテンキーからなる。また、ヘッドセットボリューム 6 4 a は、ヘッドセットから出力される音量を調節し、スピーカボリューム部 6 4 b は、スピーカから出力される音量を調節する。サウンドエフェクト選択部 6 5 は、ユーザが使用可能なサウンドエフェクト音源データファイルの表示するものであり、例えば銃声音、雷音、拍手の音、歓声音等を選択でき、サウンドエフェクト制御部 6 6 により、効果音の再生及び停止、並びに音量調節を行う。これにより、効果音の各種効果音でユーザが通話相手への気持ち等を表現することができる。

10

【0063】

また、BGM 選択表示部 6 7 は、ユーザが使用可能な例えばタヒチの波の音等の BGM 音源データファイルを表示するものであり、更に BGM 制御部 6 8 により、BGM の再生及び停止、並びに BGM の音量調節を行うことで、サウンドエフェクトと同様、ユーザ自身が選択し、調節した音量により、ユーザの気分やその場の雰囲気を通話相手へ伝えること等ができる。

【0064】

更に、音質制御部 6 9 は、音声のみのコミュニケーションを行うときと、BGM 及び効果音付きの高音質なコミュニケーションを行うときとで、ユーザが音質を選択することができ、これにより、エンコーダが圧縮符号化方式を変更し、伝送レートが可変となる。

20

【0065】

次に、ソフトウェアモジュール 1 を実行する V o I P クライアントのハードウェア構成について説明する。図 5 は、V o I P クライアントのハードウェア構成を示すブロック図である。図 5 に示すように、V o I P クライアント 2 0 において、C P U (Central Processing Unit) 2 2 1 は、R O M (Read Only Memory) 2 2 2 に記憶されている上記ソフトウェアモジュールを構成する各種プログラム、又は記憶部 2 2 8 から R A M (Random Access Memory) 2 2 3 にロードされた上述のソフトウェアモジュール 1 を構成する各種プログラムに従って各種の処理を実行する。また、図示せぬタイマが計時動作を行い、時刻情報を C P U 2 2 1 に供給する。R A M 2 2 3 にはまた、C P U 2 2 1 が各種の処理を実行する上において必要なデータ等も適宜記憶される。

30

【0066】

C P U 2 2 1、R O M 2 2 2 及び R A M 2 2 3 は、バス 2 2 4 を介して相互に接続されている。このバス 2 2 4 にはまた、入出力インターフェイス 2 2 5 も接続されている。

【0067】

入出力インターフェイス 2 2 5 には、キーボード、マウス等よりなる入力部 2 2 6、C R T、L C D 等よりなるディスプレイ、並びに、ヘッドフォンやスピーカ等よりなる出力部 2 2 7、ハードディスク等より構成される記憶部 2 2 8、モデム、ターミナルアダプタなどより構成される通信部 2 2 9 が接続されている。

40

【0068】

通信部 2 2 9 は、図示せぬインターネットを介しての通信処理を行う。C P U 2 2 1 から提供されたデータを送信する。また通信部 2 2 9 は通信相手から受信したデータを C P U 2 2 1、R A M 2 2 3、記憶部 2 2 8 に出力する。記憶部 2 2 8 は C P U 2 2 1 との間でやり取りし、情報の保存・消去を行う。通信部 2 2 9 はまた、他のクライアントとの間で、アナログ信号またはデジタル信号の通信処理を行う。

【0069】

入出力インターフェイス 2 2 5 にはまた、必要に応じてドライブ 2 3 0 が接続され、磁気ディスク 2 4 1、光ディスク 2 4 2、光磁気ディスク 2 4 3、或いは半導体メモリ 2 4

50

4等が適宜装着され、それらから読み出されたコンピュータプログラムが、必要に応じて記憶部228にインストールされる。

【0070】

次に、このようなハードウェア構成のV o I Pクライアントが上述の図2に示すソフトウェアモジュール1を構成する各種プログラムを実行することによりV o I P通話を実行する方法について説明する。

【0071】

図6は、図3に示す一般的なPC向けソフトウェアモジュールのうち、V o I P通話制御部7が制御するR T P / R T C Pパケットの送受信機能を説明する図である。図6に示すように、V o I P通話制御部7の制御は、スレッド1~スレッド5に分けることができる。スレッド1は、送信者がR T Pパケットを送信するまでの処理、即ち送信者が発話した音声を含むデータを格納したデータパケットを送信する処理であり、スレッド2、3は、送信されたR T Pパケットを受信して再生するまでの処理、即ち上記送信者が通信対象となっている相手の音声を含むデータを聞くまでの処理である。

10

【0072】

また、スレッド4は、オペレーティングシステム(O S)3に付随するスレッドライブラリにより自動生成される。スレッドライブラリは、プライオリティに応じてスレッド1、スレッド2及びスレッド3のメインプロセッサ(図5に示すC P U 2 2 1)上での計算資源配分、即ちスケジューリングを行う。

【0073】

また、スレッド5は、アプリケーションであるG U I 8のメインスレッドであり、スレッド1、スレッド2及びスレッド3を生成したり破棄したりすると共に、プログラミングされたアルゴリズム或いはユーザ操作に応じてスレッド1、スレッド2及びスレッド3の制御を行う。

20

【0074】

R T P / R T C Pパケットの送信処理であるスレッド1においては、マイクからユーザの音声をキャプチャしてP C Mデータを受け取り(Capture)、必要に応じて、B G M合成部15等により、キャプチャサンプル・フレーム毎に、効果音及びB G Mと音声とを合成し(Effect or Mixing)、コーデック13は、そのデータを圧縮符号化する(encode)。この際、コーデック13は、上述したように、他のV o I Pクライアントから受信したR T C Pパケットに含まれる制御情報の圧縮率を制御する情報に基づき、圧縮符号化方法を決定する。そして、R T P / R T C P制御部12が、圧縮符号化したデータをR T Pパケット化し、また、V o I P通話制御部7が生成した他のV o I Pクライアントが圧縮すべきエンコード帯域及びサブバンド分割数等の制御情報をR T C Pパケット化し、(packet)、これらのパケットを送信する(send)。

30

【0075】

また、R T Pパケットの送信処理とは別に、通信対象の通話装置におけるコーデックの圧縮符号化方式を制御するための制御情報を格納したR T C Pパケットの送信処理が行われる。

【0076】

一方、R T Pパケットの受信処理においては、V o I Pクライアントの受信側に設けられたデコード処理前後の揺らぎ吸収バッファ、即ち、エンコードされたバイトストリームP C Mを格納する揺らぎ吸収バッファ(Forward Jitter buffer) B F 1と、デコードされ無圧縮のlinear P C Mとされたデータを格納するP C Mバッファ(Backward Jitter buffer) B F 2とを使用する。

40

【0077】

そして、R T P / R T C Pパケットの受信機能のうち、スレッド2では、R T P / R T C P制御部12がR T Pパケットを受信し(Receive)、ネットワーク上で発生したデータ損失及び転送パケットのデータエラー等の欠落パケットを補う処理を行い(parse)、デコード処理前に設けられた揺らぎ吸収バッファB F 1に格納する(push&pop)する。

50

これをコーデック 1 3 が復号伸張し (decode)、デコード処理後に設けられる揺らぎ吸収バッファ (PCM バッファ) B F 2 に、無圧縮のデータとして格納する (push)。

【 0 0 7 8 】

そして、スレッド 3 においては、U S B カメラドライバ 2 a、U S B オーディオドライバ 2 b 及び各種ドライバ 2 c が、揺らぎ吸収バッファ B F 2 から無圧縮のデータを読み出し (pop)、この読み出したデータを再生する (sound device)。

【 0 0 7 9 】

また、スレッド 1 にて行うエフェクト効果及び B G M と音声との合成 (Effect or Mixing) は、スレッド 3 において、揺らぎ吸収バッファ B F 2 からデータを読み出した後であって、再生する前に行ってもよく、また無圧縮のデータを格納する PCM バッファを用意せず、揺らぎ吸収バッファからデータを読み出し復号したデータを揺らぎ吸収バッファ B F 2 に格納することなく再生するようにしてもよい。

10

【 0 0 8 0 】

更に R T P パケットの受信処理とは別に、V o I P 通話制御部 7 がコーデック 1 3 における圧縮符号化方式を制御するため、受信した R T C P パケットに含まれる制御情報を取り出す等の R T C P パケットの受信処理が行われる。

【 0 0 8 1 】

次に、R T P パケットに格納する実時間音響データのコーデックの変更制御方法について更に詳細に説明する。本実施の形態における V o I P システムにおいては、ネットワークの輻湊状態に応じ、送信する音声データに B G M 及び効果音等が合成された送信データの圧縮率、即ちデータの伝送レートを再呼制御することなく異ならせることにより、高音質で情報量が多い送信データであっても、通信を継続可能にするものである。

20

【 0 0 8 2 】

具体的には、例えばネットワークが何らかの要因で混雑した場合には、再呼制御を行わずにそのコーデック (圧縮符号化方式) を変更してビットレートを下げること等が挙げられる。

【 0 0 8 3 】

このため、R T P / R T C P 制御部 1 2 は、受信する R T P パケットからネットワーク伝送レートを常時又は定期的に計測する。この伝送レートの計測結果に応じて、V o I P 通話制御部 7 は、R T C P パケットにて、欠落パケットの再送要求を行なわせると同時に、通信相手が圧縮すべき周波数成分を通知させる。

30

【 0 0 8 4 】

具体的には、例えば、R T P パケットの伝送遅延 (時間的な遅れ) が大きくなる等して伝送レートが小さくなり、V o I P 通話制御部 7 がネットワークが混雑していると判断した場合、R T P / R T C P 制御部 1 2 が、送信音の圧縮すべき周波数として、例えば高周波成分を削除又はビット割当てを減らして圧縮する等するように通知するための R T C P パケットを生成するよう制御する。

【 0 0 8 5 】

これにより、例えばビットレートを下げるような指示 (制御情報) を有する R T C P パケットを受けとった他の V o I P クライアントでは、当該他の V o I P クライアントの R T P / R T C P 制御部から V o I P 通話制御部を介してこの制御情報がコーデックに入力され、コーデックが圧縮符号化方式を変更する。こうして圧縮符号化方式がコーデックにより変更され送信レートが小さくされた送信データが R T P / R T C P 制御部 1 2 から R T P パケットとされ、他の V o I P クライアントから送信される。こうして、V o I P クライアントは、通話相手となる他の V o I P クライアントから送信される R T P パケットの伝送レートを制御する制御情報を R T C P パケットにより送信し、他の V o I P クライアントから送信される R T P パケットの伝送レートを制御することができる。

40

【 0 0 8 6 】

このように、受信側となる V o I P クライアントの要求により、送信側となる他の V o I P クライアントが圧縮符号化方式を変更することにより、高音質な音声と B G M 及び効

50

果音が合成された合成音源データ、即ち高いビットレートで送信する必要があるデータ通信中に、何らかの要因でネットワーク帯域が急激に混雑し始めた場合、たとえ音質が一時的に低下したとしても会話（通信）を遮断することなく、継続させることができる。即ち、SIP等の呼制御のためのメッセージを使用することなく、RTP/RTPC/Pケットを使用して圧縮符号化方式の変更を行うことができるので通話が遮断されることがない。

【0087】

また、本実施の形態におけるRTP/RTPC/Pケットは、後述するように拡張可能なRTP/RTPC/P/AP/Pケットであり、RTP/RTPC/Pケットを使用して、欠落パケットの再送要求を行うことができ、この欠落パケットの再送要求を行うと同時に、他のVoIPクライアントが圧縮すべき周波数成分を通知することができるため、処理が迅速になる。

10

【0088】

次に、本実施の形態におけるVoIPシステムにおいて更に詳細に説明する。VoIPクライアント間における上述したRTP/RTPC/Pケットの送受信は、先ず、一方のVoIPクライアントが他方のVoIPクライアントに呼シグナリングを行い、両者が通話相手のIPアドレス、ポート番号、音声及び音響圧縮方式等を認識する必要がある。ここでは、図3に示すようなVoIPクライアントアプリケーション1を備えたPC(Personal Computer)がUA(User Agent)となり、このUA同士でVoIPする際の呼制御方法について説明する。

【0089】

図7は、図3に示すVoIPソフトウェア1を使用して通話を行う際に、やり取りされるメッセージ・データの流れを示す図である。

20

【0090】

図7に示すUA1、UA2は、SIPによってVoIPの呼制御を行うSIPUAである。このようなVoIP通信において、SIPを用いた開始側（発呼側）として動作している場合はUAC(UA Client)となり、通信の受け手側（着呼側）として動作している場合はUAS(UA Server)となり、両VoIPクライアント111、121間でデータ（メッセージ）がやり取りされる。SIPのメッセージには、UACからUASへ送られるメッセージであるリクエストと、UASからUACへ返されるメッセージであるレスポンスとがあり、リクエストのメッセージ・ボディには、SDP(Session Description Protocol)に基づいたセッション情報が収納される。また、OKレスポンス等の所定のレスポンスには、通話に使用するコーデックを選択するためのSDPメッセージを含む。

30

【0091】

先ず、発呼側のUA1から着呼側のUA2へセッション参加要求としてINVITE SDP1(D1)が送られる。着呼側のUA2にこれが届くと、着呼側のUA2からリクエストが受信され処理中であることを示す100Tryingレスポンス及び180 Ringingレスポンス(D2)が送られ、次いで、発呼側のUA1からリクエストが受理されたことを示す200 OKSDPレスポンス(D3)が送られる。この180 Ringingレスポンスは、着呼側のUA2にて呼び出し音を鳴らしている状態を示し、着呼側のUA2にて受話器を取り上げたときに、200 OKSDPレスポンスが送られる。そして、発呼側のUA1からACK(D4)が送られることにより、通信のためのポートが開かれる。こうしてRTP/RTPC/Pのセッションを確立し、音声及び音響データを送受信し合う(D5)。

40

【0092】

また、逆に着呼側であったUA2が発呼側となり、UA1へセッション参加要求INVITE SDP2(D6)が送られた場合、着呼側となったUA1から発呼側UA2(クライアント側)のエラーを伝える405 Not Acceptableレスポンス(D7)が送信された場合、これを受信した発呼側UA2は、ACK(D8)を送信し、再び着呼側のUA1へセッション参加要求INVITE SDP2(D9)を再送する。

【0093】

着呼側UA1がこれを正常に受信できれば、これを知らせる200 OKレスポンス(D10)が送られ、発呼側からACK(D11)が送られることにより、新しいRTP/RTPC/P

50

Pのセッションを確立し、音声及び音響データを送受信し合う(D12)。

【0094】

このように、先ず始めに呼制御を行ってからRTP/RTCPのセッションを確立し、音声及び音響データを送受信し合う。

【0095】

図8は、図7に示す呼制御の際に使用されるSDPの記述例を示す図である。図8に示すように、SDPで示されるタグ(注釈記号)のうち、「v」はバージョン番号を示し、「o」はセッション生成者とセッション識別用の情報を示し、o = <ユーザ名> <SDPメッセージ識別番号> <メッセージ識別サブ番号> <ネットワーク・タイプ> <アドレス・タイプ> <アドレス>と記述される。また「s」はセッションのタイトルを示し、「c」はコネクションデータを示し、s = <開始時刻> <終了時刻>と記述される。

【0096】

また、「m」はメディア名とトランスポートアドレスを示し、m = <メディア種別> <ポート番号> <トランスポート・プロトコル> <ペイロード・タイプ>と記述され、「a」はメディア属性を示し、m = rtpmap: <ペイロード・タイプ> <符号化則> / <クロック>と記述される。

【0097】

このように、SIPは、IPアドレス・ポート番号・音声及び音響圧縮方式を記述することができ、呼制御の際にはこのSDPを、お互いに送受信し合う。VoIPシステムにおいては、このSIPによりRTP/RTCPのセッションを確立し、RTP/RTCPパケットの送受信を行うが、上述した如く、音声及び音響圧縮方式をこのSIPの呼制御において確立してしまうと、ネットワーク帯域の混雑の度合いによって伝送ビットレートを可変とするために、再び再呼制御が必要となる。

【0098】

そこで、本実施の形態においては、RTCPパケットを拡張したAPPパケットを使用して、受信側にてデータの到着状態、即ち送信側の伝送レートを常時又は定期的に計測し、この伝送レートに基づき、伝送レートを上げるか、又は下げるかの判断を行い、これをAPPパケットにより送信側に通知して送信するデータの伝送レートを制御するものである。

【0099】

次に、RTP/RTCPのセッションを確立した後送受信されるRTP/RTCPパケットについて説明する。図9(a)及び(b)は、RTPパケットの夫々構成及びヘッダのフォーマットを示す図であり、図10(a)乃至(c)は、アプリケーションによって拡張可能なRTCP APPパケットの構成、そのヘッダ(RTCP APP Application-defined Header)、及び送信レポートのフォーマット例を示す図である。

【0100】

RTPは、インターネット等のIPネットワークにおいて、リアルタイムに音声や動画を送信/受信するトランスポート・プロトコルであり、RFC1889で勧告されている。RTPは、トランスポート層に位置し、一般にUDP上でRTCPと共に用いられる。

【0101】

そして、図9(a)に示すように、RTPパケットは、IPヘッダ、UDPヘッダ、RTPヘッダ及びRTPデータからなる。そして、RTPヘッダには、図9(b)に示すように、先頭から、バージョン情報格納部(V: version、例えばV = 2) F1、パディング格納部(P: padding) F2、拡張ビット格納部(X: extension) F3、CSRC (contributing source) カウント格納部(CC) F4、マーカ情報(M: marker) 格納部 F4、マーカ・ビット格納部(M: maker) F5、ペイロード種別情報格納部(PT: payload type) F6、シーケンス番号情報格納部(sequence number) F7、タイムスタンプ格納部(time stamp) F8、SSRC識別子格納部(synchronization source identifier) F9、CSRC識別子格納部 F10が設けられ、CSRC識別子格納部 F10の後ろに実時間音響データが付加される。

10

20

30

40

50

【 0 1 0 2 】

バージョン情報格納部 F 1 には、R T P のバージョンを示す情報が格納され、例えば R T P 2 を示すときには、その旨のバージョン情報が格納される。

【 0 1 0 3 】

カウント格納部 F 4 は、ヘッダ中に示される C S R C (寄与送信元識別子) の数を示す。

【 0 1 0 4 】

ペイロード種別情報格納部 F 6 には、実時間音響データの種別を示す情報が格納され、例えば映像や音声を示す旨の情報等が格納される。

【 0 1 0 5 】

シーケンス番号情報格納部 F 7 には、R T P セッションにおいて、R T P パケットを送受信する度にカウントアップされ、送受信する R T P パケットの順番を認識するためのシーケンス番号が格納される。

【 0 1 0 6 】

タイムスタンプ格納部 F 8 には、実時間音響データを作成、更新した日時に関するタイムスタンプ情報が格納される。

【 0 1 0 7 】

S S R C 識別子格納部 F 9 及び C S R C 識別子格納部 F 1 0 には、R T P セッションにおいて、データ送信側のソースを識別するための情報が格納される。S S R C 識別子は、同期送信元識別子であり、同一ユーザが組み合わせて扱うべき複数のストリームが同じ値を共有するように割り当てた識別子であり、C S R C 識別子は、寄与送信元識別子であり、ストリーム源を示す。複数のストリームがミキシング処理され1つのストリームデータとして提供される場合等に使用される。

【 0 1 0 8 】

図 3 に示す R T P / R T C P 制御部 8 は、R T P に従って実時間音響データを送信するに際して、上記各格納部に各種情報を格納すると共に、各格納部に格納された各種情報を認識して実時間音響データを抽出する処理をする。

【 0 1 0 9 】

上述の R T P が音声・動画像データそのものを送信/受信するプロトコルであるのに対し、R T C P は、周期的に、パケットロス、遅延ジッタ、ラウンドトリップ等の回線品質を評価し、その帯域に見合ったリアルタイム通信を実現するため情報を送信/受信するプロトコルである。

【 0 1 1 0 】

この R T C P を用いることにより、相手からフィードバックされてくる情報により、ネットワークの状態などを推測して送信レートを変更するなどの動的な処理を行うことができる。また、今誰がデータを送信していて、誰が受信しているかを示す情報も R T C P パケットで同時に送っているので、今現在の参加者の情報を知ることできる。

【 0 1 1 1 】

図 1 0 (a) に示すように、R T C P パケットは、I P ヘッダ、U D P ヘッダ、R T C P ヘッダ及び R T C P データからなる。そして、拡張可能な R T C P パケットである R T C P A P P パケットのヘッダには、図 1 0 (b) に示すように、先頭から、バージョン情報格納部 (V : version、例えば V = 2) F 1 1、パディング格納部 (P : padding) F 1 2、subtype 格納部 F 1 3、パケットタイプ (P T : packet type) 格納部 F 1 4、レポート長格納部 (length) F 1 6、S S R C / C S R C 識別子格納部 F 1 7、アスキー (A S C I I : American Standard Code for Information、情報交換用アメリカ標準コード) で記述される N a m e 格納部 F 1 8 が設けられ、この後に、アプリケーション独自のデータが格納されるデータ格納部 (Application-Dependent Data) F 1 9 が付加される。

【 0 1 1 2 】

パケットタイプ P T 格納部 F 1 4 には、R T C P パケットの種別が記録され、本実施の形態においては、このパケットタイプ P T = A P P (Application : アプリケーション固

10

20

30

40

50

有情報) = 204 (パケットタイプ値) と記述される。APPは、RTP規定外のアプリケーション固有の制御情報を通知するためのパケットであることを示す。

【0113】

図3に示すRTP/RTP制御部8は、RTP APPパケットとして、上述したように、欠落パケットの再送要求、及びRTPパケットを受信した際のジッタを吸収するために使用するRTPパケット受信のための揺らぎ吸収バッファのサイズ等をデータ格納部F19に記述して送信することができる。そして同時に、VoIP通話制御部3の指示により、通信相手が送信すべき送信データにおける圧縮する周波数成分の情報を記述して送信することができる。

【0114】

ここで、RTPパケットには、RTPデータの送信者から送られるタイプのRTP SR (Sender Report) パケットと、RTPデータの受信者から送られるタイプのRTPパケットRTP RR (Receiver Report) パケットとがある。

【0115】

SRパケットは、ストリームを送出している端末から他の端末に対して送られるもので、自装置が送出したストリームに関する情報である送信情報 (sender info) と、受信したストリーム各々についてストリームの受信状態 (パケット破棄率、ジッタ等) を送信装置へ報告するためのレポートブロック (reception report block) とを含み、RRパケットは、他の通話装置から受信したスリームに関する情報を通知するためのもので、同じく受信したストリーム各々についてストリームの受信状態を送信装置へ報告するためのレポートブロックを含むものである。

【0116】

このレポートブロックは、図10(c)に示すように、パケットの送信者の同期送信元 (SSRC: Synchronization Source) 識別子、RTP損失率、損失RTPパケット数、受信シーケンス番号、到着時間間隔のジッタの平均値、最後に受信したSRの送信時刻 (LSR: Last SR timestamp)、最後にSRを受信した時刻からこのRRを送るまでの時間 (DLR: Delay since Last SR) を入れることになっている。

【0117】

したがって、送信側においては、RTPデータの送信の際に、送信RTPパケット数及び送信RTPバイト数を管理しておき、また、受信側においては、RTPデータの受信の際に、受信RTPパケット数、損失RTPパケット数及び到着時間のジッタ等の管理情報を管理する。

【0118】

図11は、本実施の形態におけるRTP APPパケットの交換時にやりとりされるメッセージを示す図である。RTPのSenderでもありReceiverでもあるUA1とUA2との間でやり取りされるSRブロックを含むRTPパケット及びRRブロックを含むRTPパケットにおいては、図11に示すように、拡張データとして、エンコード帯域幅 (Encode Bandwidth)、サブバンド分割ブロック数 (sub band Numbers)、デコード処理部前後の揺らぎ吸収バッファサイズ (時間に換算可能) (Forward and Backward Jitter Buffer size)、上記吸収バッファにキューイングされているデータサイズ (時間に換算可能) (Buffer queued size)、RTPパケットの再送リクエスト情報 (Re-Request: sequence number) を記述することができる。RTPパケットの再送リクエストにはシーケンス番号を利用する。

【0119】

UA1は、UA1のユーザの音声にBGM・効果音等が合成された送信音源ストリームをRTPパケットにて送信する。また、UA2も、UA2のユーザの音声にBGM・効果音等が合成された合成音をRTPパケットにて送信する。

【0120】

そして、例えばUA1は、RTP APPパケットに、ストリーム受信装置としてのUA1がストリーム送信装置としてのUA2に対してそのストリーム受信状態を報告する

10

20

30

40

50

レポートブロックに加え、U A 2 から送られる R T P パケットの到着状態に応じて、制御情報としてのエンコード帯域幅及びサブバンド分割ブロック数を記述する。これを受けた U A 2 は、そのエンコード帯域幅及びサブバンド分割ブロック数に基づき、U A 1 に送信する送信データのエンコード帯域及びサブバンド分割ブロック数を変更することで、伝送レートを変更する。

【 0 1 2 1 】

また同様に、U A 2 から U A 1 に対して、U A 1 から送られる R T P パケットの到着状態に応じて伝送レートを制御するための制御情報としてエンコード帯域幅及びサブバンド分割ブロック数を通知することができ、同様に、これを受けた U A 1 はエンコードすべき帯域及びサブバンド分割ブロック数を変更し、伝送レートを変更する。

10

【 0 1 2 2 】

更に、U A 1 と U A 2 との間にて、S R (Sender Report) ブロックを含む R T C P パケットを送信してから S R (Receiver Report) ブロックを含む R T C P パケットを受信するまでの時間が R T T (Round Trip Time) として求められ、U A 1 は、この R T T と、自身のデコード処理部前後の揺らぎ吸収バッファサイズとに応じて U A 2 から再送してもらうことが可能な欠落パケットのシーケンス番号を通知する。即ち、R T T の計算結果に基づき、再送要求したパケットが送られてくるまでの時間情報が得られ、再送要求したパケットが到着する時刻よりそのパケットのデコードを開始予定時間又は再生予定時間が遅いときは再送要求を行うことができる。

【 0 1 2 3 】

なお、このアプリケーション固有情報 A P P として、自身のデコード処理前後の揺らぎ吸収バッファサイズ、この揺らぎ吸収バッファを占有しているデータ量等の情報も記載することができ、これらの通知を受けた通話相手側の V o I P クライアントは、これらの値に基づき、送信するデータの伝送レートを変更する等してもよい。

20

【 0 1 2 4 】

このように、U A 1 及び U A 2 において、R T C P A P P パケットに含まれるエンコード帯域幅及びサブバンド分割ブロック数により、通信相手から送られる R T P パケットの圧縮すべき音の周波数成分を、通話相手に通知することができ、高音質な音声通話中に、何らかの要因でネットワーク帯域が急激に混雑しはじめた場合であっても、伝送レートを下げることにより、会話を遮断されることなく、継続させることができる。

30

【 0 1 2 5 】

即ち、本実施の形態においては、例えば U A 1 からの送信データが格納された R T P パケットを受信した受信側の U A 2 は、その受信結果に基づき U A 1 に伝送レートの変更を要求する R T C P A P P パケットを送信する。これを受け取った U A 1 が、受け取った A P P パケットに含まれる制御情報に基づき、U A 2 に対して欠落パケットを再送すると共に、圧縮する周波数成分の情報に基づき送信データの圧縮符号化方式を変更し、よって伝送するデータの伝送レートを可変とすることができる。このように、R T P パケットの受信側の要求により、R T P パケットを送信する送信側の V o I P クライアントがコーデックを適宜変更したデータを伝送することにより、コーデックの変更通知をするための再呼制御が不要になり、通信が途切れることがない。

40

【 0 1 2 6 】

更に、本実施の形態においては、常時又は定期的に受信側において伝送レートを計測し、その伝送レートに応じて送信側のコーデックを自動的に制御するものとしたが、受信端末を使用しているユーザ操作に基づき同様に処理を行うことができる。

【 0 1 2 7 】

即ち、図 4 に示す G U I 8 に示したように、例えば、伝送レートを上下するための音質制御部 6 9 等の操作部を設けておき、音声のみの通話か又は音声に B G M 等を付加した高音質のデータによる通話かに応じてユーザ自身が伝送レートを指示できるようにしたり、又はユーザが例えば音声及び音響が聞き取りにくいと感じた場合には、ユーザ自身がこの操作部を操作して伝送レートを変更することができる。また、伝送レートを下げる際に高

50

周波成分をカットするか、音響データをカットするか等を選択できるような操作部を設け、ユーザ操作に基づき V o I P 通話制御部 3 が R T C P A P P パケットを使用して、送信者側の伝送レートを制御するようにしてもよい。また、逆に、ユーザの要望に応じて伝送レートを上げて音質を向上可能なようにしておいてもよい。

【 0 1 2 8 】

次に、パケットを送信してからメッセージの往復に要する時間を示す R T T の求め方について説明する。この R T T は、上述したように、V o I P 通話制御部 3 によち、欠落パケットの再送要求を行う際、R T P パケットを受信した際のジッタを吸収するために使用する R T P パケット受信バッファのサイズと共に参照されるものである。

【 0 1 2 9 】

図 1 2 は、R T C P を利用した R T T の計算方法を示すシーケンス図である。図 1 2 に示すように、R T P の Sender でもあり Receiver でもある U A 1 と U A 2 との間にて、まず、U A 1 から R T C P S R パケットが送られると、U A 2 から R T C P R R パケットが送られる。R T C P R R パケットには、レポートブロックとして L S R 及び D L S R が記述されている。

【 0 1 3 0 】

L S R は、最後に受信した S R パケットの N T P (Network Time Protocol) タイムスタンプであり、例えば図 1 2 においては、U A 1 が R T C P S R パケットを送信した時間 (10 Nov 1995 11:33:25.125) である。R R パケットを受信した U A 1 は、R T C P R R パケットの到着時間 A (10 Nov 1995 11:33:36.5) と、R R パケットのレポートブロックに記載された L S R とから R T T (= L S R - A) を算出することができる。

【 0 1 3 1 】

R T T には、最後に S R パケットを受信したとき、このレポートブロックを送信するまでの経過時間を 1 / 6 5 5 3 6 秒単位で表した値を示す D L S R 等が含まれる。これは、伝送レートやネットワークの混雑度合いによって変化するものである。

【 0 1 3 2 】

こうして U A 1 は、算出した R T T 及びパケット到着時刻等によりネットワーク負荷を推定し、欠落パケットの再送要求をしたり、ストリームの伝送レートを調整する。この R T T は、R T C P パケットを送受信する毎に算出してもよく、又は所定時間経過毎に算出したり、コーデックを変更した際に算出するようにしておよい。

【 0 1 3 3 】

次に、聴覚心理特性を利用した高能率音響圧縮符号化及び復号伸張方法の一具体例について説明する。図 1 3 及び図 1 4 は、夫々高能率音響圧縮符号化部及び高能率音響復号伸張部を示すブロック図である。高能率音響圧縮符号化部 7 0 は、図 2 に示すエンコーダ 3 1 に対応するものであり、図 1 3 に示すように、時間周波数分解部 7 1 と、量子化部 7 2 と、聴覚心理モデル部 7 3 と、帯域割当て部 7 4 と、マルチプレクス 7 5 とを有する。

【 0 1 3 4 】

時間周波数分解部 7 1 は、時間軸の信号を所定時間単位でブロック化あるいはフレーム化してこのフレーム毎の時間軸の信号を周波数軸上の信号に変換 (スペクトル変換) して複数の周波数帯域に分割する。

【 0 1 3 5 】

聴覚心理モデル部 7 3 は、一般に高域ほど帯域幅が広くなるような臨界帯域 (クリティカルバンド) と呼ばれている帯域幅で、オーディオ信号を複数 (例えば 2 5 バント) の帯域に分割する等の処理を行い、帯域割当て部 7 4 は、各帯域毎に所定のビット配分或いは、各帯域毎に適応的なビット割当て (ビットアロケーション) を行う。例えば、モディファイド離散コサイン変換 (modified discrete cosine transform : M D C T) 処理されて得られた係数データをビットアロケーションによって符号化する際には、各フレーム毎の M D C T 処理により得られる各帯域毎の M D C T 係数データに対して、適応的なビット数が割当てられる。

【 0 1 3 6 】

10

20

30

40

50

ここで、上述した R T C P パケットに格納される制御情報に含まれるエンコード帯域及びサブバンド分割ブロック数は、例えば聴覚心理モデル部 7 3 に供給され、エンコードすべきデータを指定されたエンコード帯域のデータのみとし、このエンコード帯域をサブバンド分割ブロック数に分割する等の処理を行う。

【 0 1 3 7 】

量子化部 7 2 は、帯域毎に割り当てられたビット数に基づき、量子化ステップ又は量子化サイズを決定して量子化する。

【 0 1 3 8 】

マルチプレクス 7 5 は、量子化されたデータを、帯域割当て部により割り当てられたビット数等の副情報と共に多重化して出力する。

【 0 1 3 9 】

また、図 1 4 に示すように、高能率音響復号伸張部 8 0 は、デマルチプレクス 8 1 と、逆量子化 8 2 と、時間周波数再構成部 8 3 とを有する。デマルチプレクス 8 1 は、高能率符号化されたデータが入力され、これをデマルチプレクスする。逆量子化部 8 2 は、出マルチプレクス 8 1 より取り出された帯域情報等の副情報に基づき量子化データを逆量子化し、時間周波数再構成部 8 3 は、時間軸上のデータを周波数軸上のデータに変換して出力する。

【 0 1 4 0 】

このような高能率音響圧縮符号化によれば、音響情報チャンネルの全てのチャンネルを合わせた総ビットレートを可変とし、かつ一定の最大値を越えないようなビット配分を行なわせることができる。

【 0 1 4 1 】

即ち、A P P パケットにて指定されたエンコード帯域により、圧縮すべき周波数成分を選択し、A P P パケットにて指定されたサブバンド分割ブロック数にエンコード帯域を分割するので、通信相手からの A P P パケットに含まれるエンコード帯域及びサブバンド分割ブロック数に基づき、出力するデータのビット数を削減したり、又は増大させたりしてエンコードする実時間音響データの圧縮率を変更することで、伝送するビットレートを可変とすることができる。

【 0 1 4 2 】

また、ここでは、A P P パケットにて圧縮すべき周波数成分を指定するものとしたが、例えば伝送ビットレートを指定するようにし、送信側の高能率圧縮符号化部 7 0 が、その指定された伝送ビットレートになるよう圧縮符号化するようにしてもよい。更に、制御情報として各帯域に割り当てるビット数や、量子化部が量子化する量子化ステップ又はサイズ等の情報を指定してもよい。

【 図面の簡単な説明 】

【 0 1 4 3 】

【 図 1 】 本発明の実施の形態における V o I P システムの一例を示す模式図である。

【 図 2 】 本発明の実施の形態における V o I P システムのうち、V o I P クライアント 1 1 1 とインターネット 1 3 0 とからなる要部を示すブロック図である。

【 図 3 】 本発明の実施の形態における V o I P システムが有する V o I P クライアントとして、一般的な P C を使用した場合における V o I P クライアントアプリケーションを示す図である。

【 図 4 】 上記 V o I P クライアントアプリケーションにおける G U I の一例を示す模式図である。

【 図 5 】 本発明の実施の形態における V o I P システムにおける V o I P クライアント (V o I P クライアント) のハードウェア構成を示すブロック図である。

【 図 6 】 図 3 に示す V o I P クライアントアプリケーションのうち、R T P / R T C P パケットを送受信する処理を示す機能ブロック図である。

【 図 7 】 上記 V o I P クライアントアプリケーションを使用して V o I P を行う際の呼制御においてやり取りされるメッセージ・データの流れを示す図である。

10

20

30

40

50

【図8】 呼制御の際に使用されるSDPの記述例を示す図である。

【図9】 (a) 及び (b) は、夫々RTPパケットの構成及びヘッダのフォーマットを示す図である。

【図10】 (a) 乃至 (c) は、アプリケーションによって拡張可能なRTCP APPパケットの構成、そのヘッダ (RTCP APP Application-defined Header)、及び送信レポートのフォーマット例を示す図である。

【図11】 本実施の形態におけるVoIPシステムにおいて、RTCP APPメッセージの交換時にやりとりされるメッセージを示す図である。

【図12】 RTPを利用したRound Trip Timeの計算方法を示すシーケンス図である。

【図13】 聴覚心理特性を利用した高能率音響圧縮符号化部を示すブロック図である。

10

【図14】 聴覚心理特性を利用した高能率音響復号伸張部を示すブロック図である。

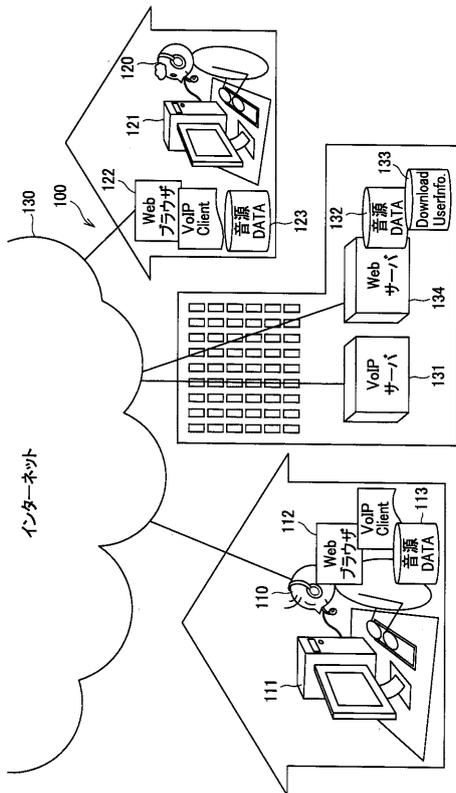
【符号の説明】

【0144】

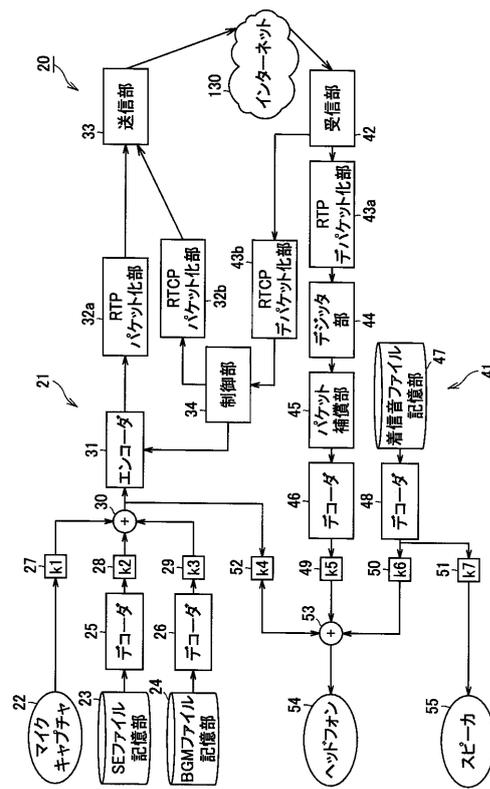
- 1 VoIPクライアントアプリケーション、21 送信手段、22 マイクキャプチャ、23 効果音ファイル読み込み部、24 BGMファイル読み込み部、25、26、31 デコーダ、27、28、29 ゲイン調整部、30 合成部、31 エンコーダ、32 パケット化部、33 送信部、41 受信手段、42 受信部、43 デパケット化部、44 デジッタ部、45 パケット補償部、46 デコーダ、47 読み出し部、48 復号部、49、50、51、52 ゲイン調整部、53 合成部、54 出力部、55 スピーカ、100 VoIP通信システム、111、121 VoIPクライアント、110、120 ユーザ、112、122 ウェブブラウザ、134 ウェブサーバ、113、123 音源データ、130 インターネット

20

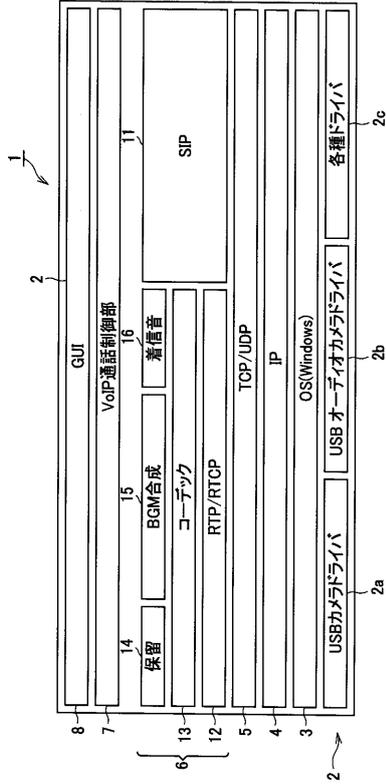
【図1】



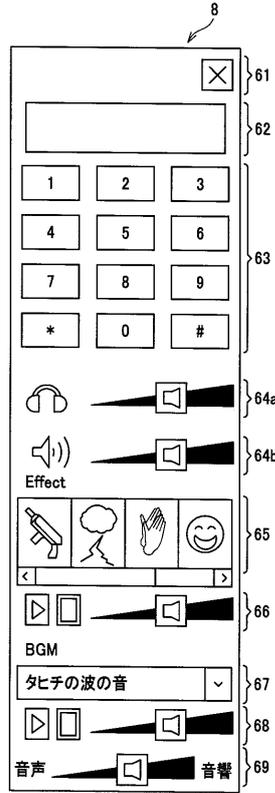
【図2】



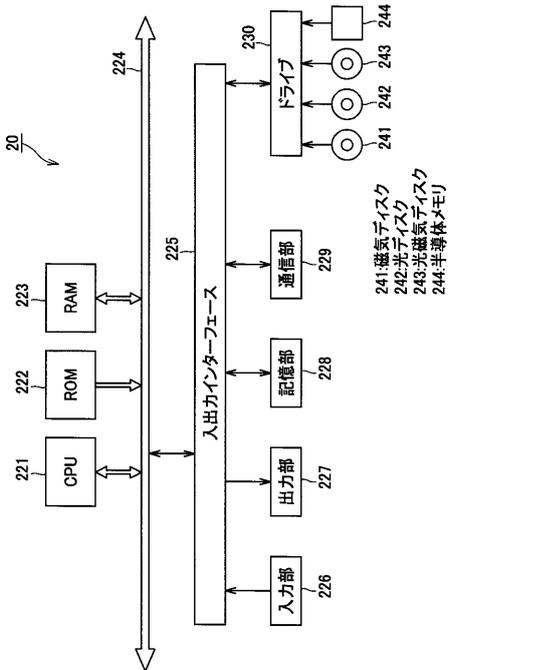
【 図 3 】



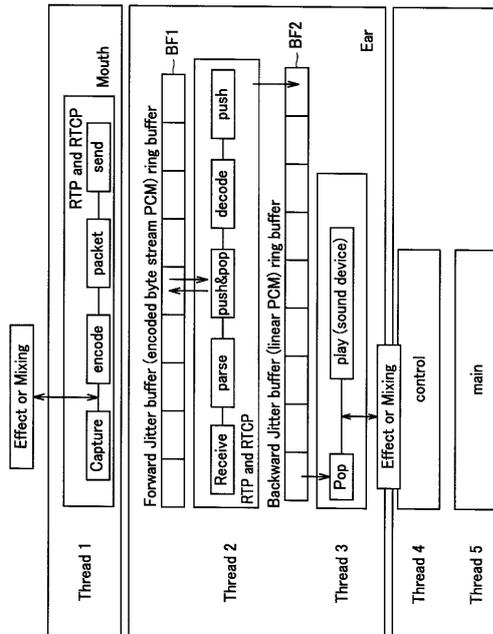
【 図 4 】



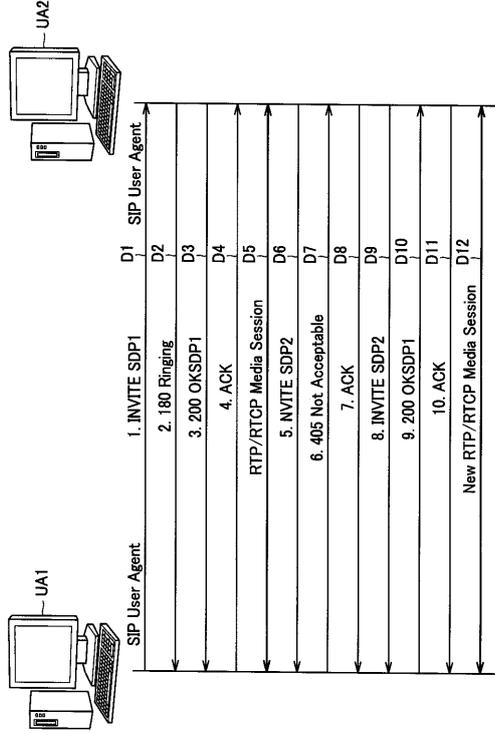
【 図 5 】



【 図 6 】



【 図 7 】



【 図 8 】

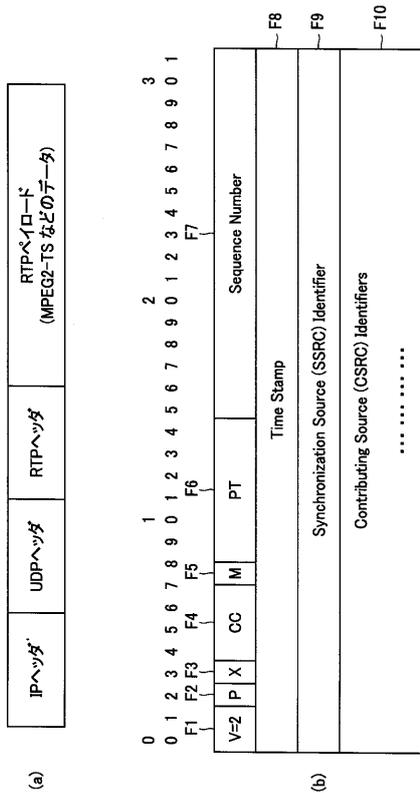
SDP Example

```

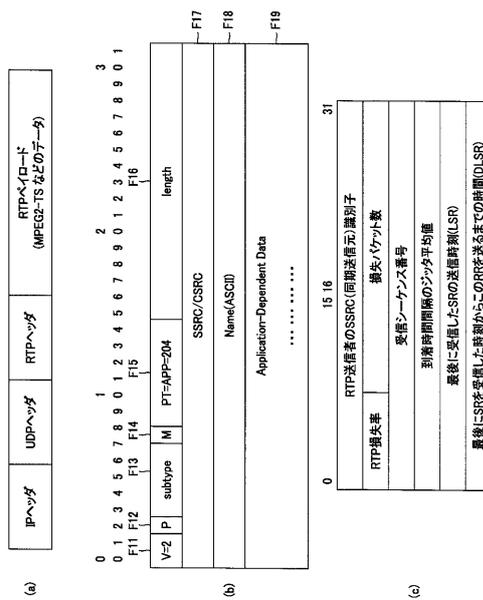
v=0
o=i
s=i
c=IN IP4 128.2.3.1
t=
m=video 4004 RTP/AVP 14 26
a=rtpmap:14 MPA/90000
a=rtpmap:26 JBEG/90000
m=audio 4006 RTP/AVP 0 4
a=rtpmap:0 PCMU/80000
a=rtpmap:4 GSM/80000

```

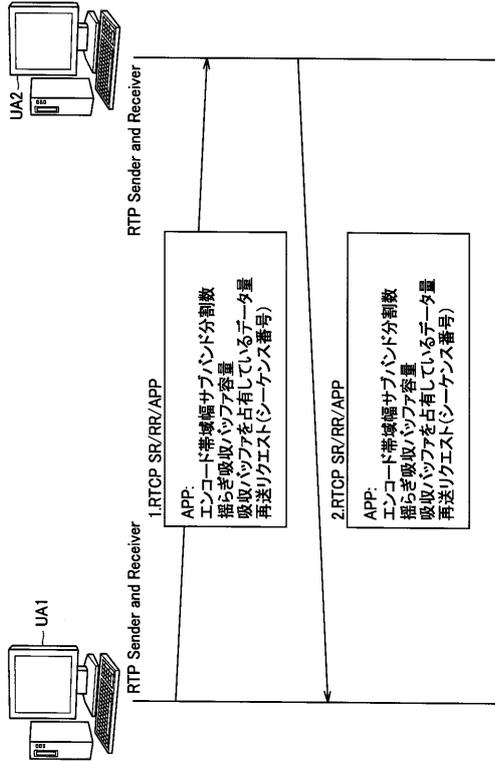
【 図 9 】



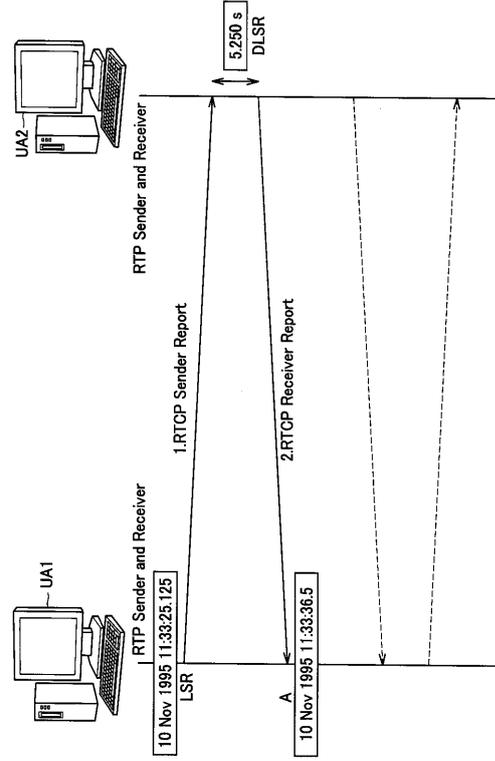
【 図 10 】



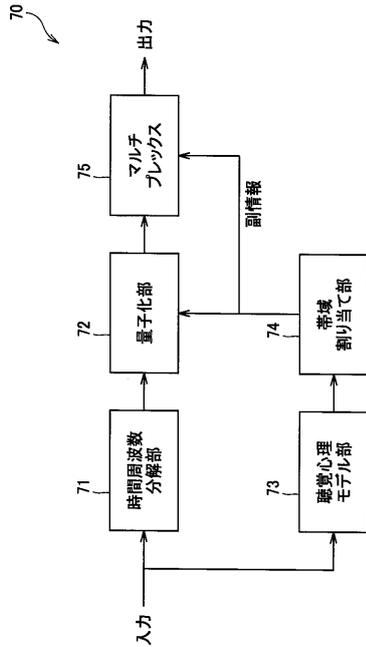
【 図 1 1 】



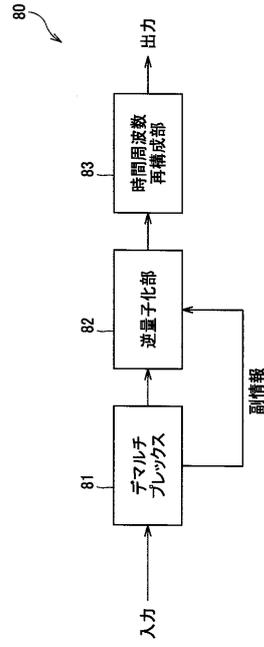
【 図 1 2 】



【 図 1 3 】



【 図 1 4 】



フロントページの続き

(72)発明者 保木本 晃弘

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72)発明者 川畑 哲

東京都品川区北品川6丁目7番35号 ソニー株式会社内

Fターム(参考) 5J064 AA02 BA16 BC11 BC18

5K030 HA08 HB01 JA05 KA03 KA19 LA01 LC01

5K101 KK20 SS07 SS08 UU19

【要約の続き】