

IP ネットワーク上のリアルタイム音声ミキシングに対して バッファサイズが与える影響に関する一考察

清 末 悌 之[†] 湯 田 佳 文^{††}

計算機ネットワーク、いわゆる IP プロトコルを用いたネットワーク上での、様々なデータの送受信が普及している。特に音声データを送受信するアプリケーションも現れている。しかし、それらアプリケーションは 1 対 1 で行う会話が主で、複数発話者が同時に発話し第三者が受信して聴取することは検討されていない。本論文では、複数発話者が存在するときに、音声品質に受信バッファのサイズ（数と単体のサイズ）が関与すると想定し、シミュレーションでその影響を示すとともに、実測して予想の正しさを示した。2 カ所から間断なく音声を送信され、着信側でミキシングする場合は、定期的に音声の途切れや抜けが生じることを示し、その間隔がバッファサイズとミキシング周期の関数で示されることを示した。

A Study of the Buffering Size Effecting to Real Time Sound Transfer on IP Network during Received Sounds Mixture

YASUYUKI KIYOSUE[†] and YOSIFUMI YUDA^{††}

In this paper we studied sound transfer process over IP network. The case that mixing process exists in receiving terminal is discussed especially. When many terminal send at the same time, receiving sounds suffer bad effects; sound skip, sound aperture, delay perturbation. We simulate sound sending-receiving-mixing process under two conditions; 1) mixing cycle is longer than buffer size, 2) mixing cycle is shorter than buffer size and clarified the mechanism causing the bad effects. We also measured delay time actually so that showed the simulations are reasonable.

1. はじめに

インターネットをはじめとする計算機ネットワークが普及している。いわゆる、IP プロトコル上に構築されたネットワーク上で、音声データを送受信することは、VoIP (Voice over IP) と呼ばれ、インターネット電話や H.323 などの具体的なアプリケーションによって、すでに珍しいものではなくなっている¹⁾。

IP ネットワーク上で音声を送信する場合、音声品質を維持するためには、まず、遅延をなるべく生じさせないようにすることが必要で、リアルタイム性を保持して音の送受信を行う研究開発はさかんである。

デジタルデータとしての音の送受信については、送信側の端末が有するバッファと、受信側の端末が有するバッファを用いたモデルで説明することができる。

基本的には、これらのバッファサイズの総和が大きいほど遅延は大きくなる。バッファサイズを動的に変化させることで、ネットワーク遅延の変動を吸収する試み²⁾などがある。

しかし、同時に送信する側が 2 カ所以上あり、受信側でミキシング処理して出力する必要がある場合については検討されていない。これは、1 対 1 の電話機能ではなく、多人数参加型のアプリケーション³⁾が台頭してきている現在、必要な技術といえる。本論文では、IP ネットワーク上で複数同時発話者が存在するときの受信側での出力音の品質について検討したので報告する。

以下、本論文では、2 章で IP ネットワークを用いたデジタル音声の送受信のモデルと、音声品質を低下させる 3 つの要素について説明し、3 章でこれら劣化要素を内包するメカニズムをシミュレーションで明らかにし、4 章で実際に計測を行ってメカニズムが正しいことを証明する。さらに 5 章では劣化要素の影響を少なくするための条件について考察する。

[†] NTT サイバースペース研究所
NTT CyberSpace Laboratories

^{††} NTT 西日本
NTT West

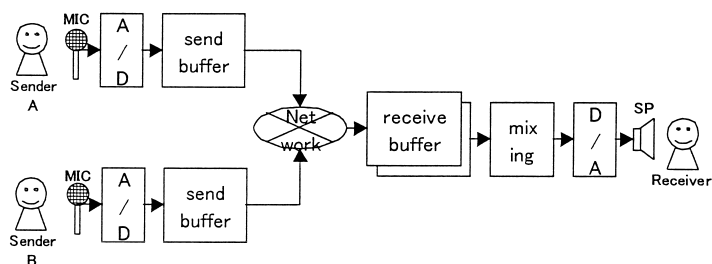


図 1 デジタル音声送受信モデル

Fig.1 Digitalized voice transfer model.

2. デジタル音声送受信のモデルと音声品質劣化の要素

2.1 デジタル音声送受信のモデル

本論文で用いる、計算機ネットワークを用いてデジタル音声を送受信するシステムのモデルを図 1 に示す。ここでは、ネットワーク内の音質劣化の要因を考慮しないことにし、図ではネットワークをブラックボックスとして示している。また、他のバッファの存在など実際はもっと複雑な構造をしているが、ここでは、以後の説明で使用する本質的な構造のみを示すにとどめている。

音声は、まず、発話側端末の利用者 (Sender) の発話を契機にして生じ、これがマイク (MIC) などの入力装置から入力される。入力された音声はデジタル化され (A/D 変換)、送信バッファ (send buffer) に蓄えられる。以後は送信処理が送信バッファの状態を見てネットワーク (Network) に送出する。これらの処理が、それぞれの発話源で行われる。図 1 では、2 カ所 (SenderA と SenderB) に送信源がある場合について示している。

受信側では、ネットワークから送りこまれるデータを、送信元それぞれに用意された受信バッファ (receive buffer) にまず蓄える。次にこの中のデータをミキシングした後、アナログ変換 (D/A 変換) してスピーカ (SP) などの出力装置を用いて出力する。これを受信者 (Receiver) が音として受け取る。ミキシングは受信バッファの状態に基づいて行われ、1 カ所からしか受信していない場合は特に処理をせずにそのまま出力する。また、1 カ所からの音声を出力中に別の着信音声があった場合、最初の出力が終了するまで (発話が終わるまで) 出力を待たされることになる。

それぞれのバッファは、処理の連続性を維持するために、2 つ以上用意され、片方の出力が行われているときも、もう片方が入力を受け付けることができるようにするダブルバッファが採用されているのが普通で

ある。

伝送遅延は CPU の処理性能が向上した現在では、これらバッファが主な原因になっている。CPU はむしろ、バッファに蓄えることのできる音声の出力時間よりも早く処理をすることができる。たとえば、発話時間 10 秒のデータをそれより短い時間で送信・伝送・受信することができる。したがって、最終的に音声出力装置に出力するときがネックとなって (10 秒の発話内容を 5 秒で出力することはできないため)、1 カ所からのみ受信する場合はよいが、複数箇所からの送信があり、受信側でミキシングをする必要がある場合は、タイミングによって、強制的に片方の出力が待たされることがあり、これが遅延の原因となる。また、本論文が議論しようとしている音声品質の劣化の原因となる部分であり、このメカニズムについては、3 章で詳しく述べる。

2.2 音声品質劣化の要素

前節でも述べたように、伝送の遅延はバッファが存在するためであり、リアルタイム性を保持するために、強制的にバッファ内のデータを出力させる手段などが考えられている⁴⁾が、遅延は原理的には避けられないものである。

本論文では、遅延を除いた、以下の 3 つの音声品質の劣化要因について考える。

- (1) 遅延の変動 (perturbation)
- (2) 音の途切れ (skip)
- (3) 音の空隙 (aperture)

遅延の変動とは、遅延時間が着信側の時間を基準にした場合、変動して出力されることであり、変動しない場合は、聴取だけして対話をしない場合は遅延していることが認識できないが、変動するために遅延の存在を認識できるというものである。対話を行っている場合は、固定的な遅延については人間は慣れることができるが、変動する場合は応答のタイミングを図りかねてストレスを生じる原因になる。

ここで、音の途切れと空隙は混乱しやすいので図を

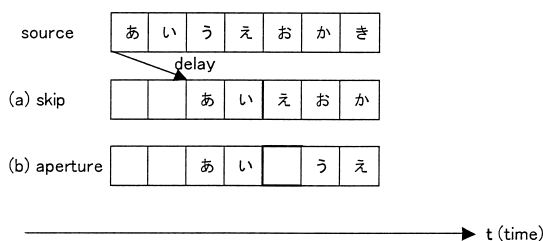


図2 音の途切れと空隙. (a) 途切れ (b) 空隙

Fig. 2 Voice skip and aperture. (a) skip (b) aperture.

用いて説明する。途切れとは、連続して出力されるべき音声の一部が消失して出力音の途中に落差が生じることをいう（図2(a)）。一方、音の空隙とは、本来連続して出力されるべき音声の一部に不必要な無音状態が生じることをいう（図2(b)）。途切れも空隙も生じなければ遅延の変動は生じないため、(1)と(2)、(3)は互いに関連性を持っているといえる。

次章では、これら音声品質に影響を与える要因の発生原因についてモデルを用いてシミュレーションを行う。

3. 劣化要因発生のシミュレーション

本章では、デジタル音声伝送を行って受信側でミキシング後に出力する場合、劣化要因の発生が避けられないことをシミュレーションによって示す。

まず、記号の定義から行う。図1で示した、送信バッファサイズを T_s とし、受信バッファのサイズを T_r とする。ここで、バッファのサイズとは、バッファに蓄えることができる音データを出力したときに要する時間のことである。すなわち、サイズ T_{sec} のバッファとは、バッファいっぱいにデジタル化した音データを蓄積した場合、これを出力装置から出力すると、 T_{sec} かかるということを意味する。以下の説明では、理解を容易にするために、 $T_s = T_r = T$ とする。一方、ミキシングの周期を μ で表すことにする。

これらの記号を用いて、以下シミュレーションを行うが、計算機ネットワークの伝送遅延や、D/A変換やミキシングなどの処理時間は無視できる、もしくは、定常的な効果を持ち、音声データの内容によって処理時間が変化しないものと仮定して説明している。この仮定は、あるバッファから別のバッファへのデータの移動が瞬時に行われること、もしくは、つねに一定時間で行われることを意味している。また、バッファ内にはつねに音データが満たされているとする。

ここでのシミュレーションは、以下の2つの条件で行う。

(1) ミキシング周期 μ がバッファサイズ T より長い場合 ($\mu > T$)

これは、ミキシングの時間間隔が出力音の時間より長い場合で、音が鳴り終わっても次のミキシングが行われていない（終了していない）パターンである。ここでは、このパターンを出力枯渇モデル（exhaustion）と呼ぶ。

(2) ミキシング周期 μ がバッファサイズ T より短い場合 ($\mu < T$)

これは、音が鳴り終わるより前にミキシングを行うために、溢れてしまうパターンである。ここでは、このパターンを出力過剰（flood）モデルと呼ぶ。

これら以外の条件として、ミキシング周期 μ とバッファサイズ T が等しい場合 ($\mu = T$) が考えられるが、これは、上記2条件の間のクリティカルなパターンで結果は自明⁵⁾なので、シミュレーションは行わなかった。

3.1 出力枯渇モデルにおけるシミュレーション

このモデルでは、ミキシングを行おうとするときに、ミキシング間隔 μ 以下の時間分の T しか音データが存在しないため、ミキシング後に鳴らし終わっても次のミキシング契機にまで至っていないということになる。したがって、ミキシング時間 μ の時間間隔で、つねに、ミキシング時間とバッファサイズ T の差分、

$$\mu - T \quad (1)$$

分の時間の空隙（aperture）が発生する。これは1回のミキシングで蓄積される遅延時間となる。また、遅延時間の累積がバッファサイズと等しくなった、

$$\mu T / (\mu - T) \quad (2)$$

時間後に、有限時間サイズのバッファに蓄積された音データをミキシング処理が受けきれなくなり、音の途切れ（skip）が生じる。すなわち、ミキシング回数を $n = 1, 2, \dots$ で示すと、

$$n\mu \geq \mu T / (\mu - T) \quad (3)$$

$$n \geq T / (\mu - T) \quad (3)'$$

で示す、 n 回ごとに、式(1)の空隙後に T 時間分の音の途切れが生じる。

遅延は、オフセット分の T に、ミキシングごとに増える式(1)の空隙分が加わり、 n 回のミキシング後に、

$$T + \alpha + (\mu - T)(n - 1) \quad (4)$$

の遅延が存在する。 α は $0 < \alpha < \mu$ の範囲にある値で、ミキシングのタイミングとバッファからの送回のタイミングのずれの時間であって、送信端末ごとに異なる。式(4)で表される遅延は T から式(1)ずつ徐々に増加し、式(3)'で示される n 回のミキシングごとに再び T に戻り、以後、 T と $2T$ の間で周期的に、増

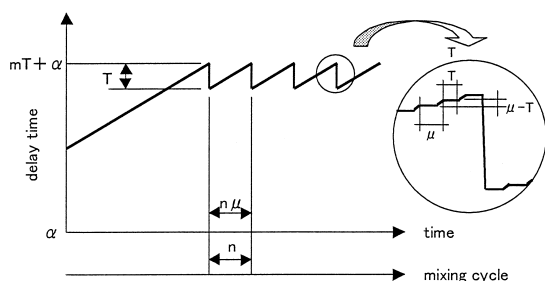


図3 出力枯渇モデルにおける遅延時間の推移

Fig. 3 Change of delay time on "exhausted" model.

加-急激減少のパターンを繰り返す。

以上はバッファが2個の場合について述べたが3つ以上の場合(ここでは、個数を m とする)では、式(2)が、

$$\mu(m-1)T/(\mu-T) \quad (5)$$

となり、式(3)'が、

$$n \geq (m-1)T/(\mu-T) \quad (6)$$

となる。また、式(4)は、

$$(m-1)T + \alpha + (\mu-T)(n-1) \quad (7)$$

となる。

以上を模式的に示すと、図3のようになる。ここで、音の空隙が生じたときは、厳密に言えば比較対象となる音がないわけであるから、遅延時間は定義上算出できないわけであるので、プロット点が存在しないものとしてグラフを示している。したがって、ミクロに示せば、 μ 時間ごとに $(\mu-T)$ 時間の間は $(\mu-T)$ 時間分の増加になっている。図3では、このことを丸の中に示した。

枯渇モデルのシミュレーション結果を図4に示す。ここでは、バッファサイズ $T = 100$ msec、ミキシング周期 $\mu = 110$ msec として示した。図中、最小単位(最小長方形の横幅)は10 msec である。また、ミキシング開始のタイミングと受信完了のタイミングを合わせるのは困難であるので、図には、1つのタイミングのときだけについて書いている。

図中、受信バッファは便宜上時間軸方向に幅を持って示されているが、伝送のことだけ考えると、ここでの仮定に基づけば伝送時間は0としているので、一瞬にして終了することになるので、図のバッファの右端の時点で送受信が行われる。例をあげると、図中、Pa1と示された送信バッファ内のデータは、100 msec をかけて伝送されるのではなく、100 msec 経過時に一瞬にして送信され、受信バッファに入るものとする。しかし、送信バッファには、実際に人間が喋ったり音を鳴らしたりするのであるから、100 msec の時間を

かけないと入力完了しない。したがって、送信、伝送、受信完了までのトータルを考えると100 msec かかるというわけである。これは1送信端末についていえることであり、複数の送信端末(図中では2台)ある場合においては、各々の端末の送受信についてこの制約条件が与えられる。

一方、受信バッファに入った後の出力では、やはり、出力装置を介して、実時間で出力し人間が聴くわけであるので、バッファに蓄えられたデータ時間分の出力時間を要することになる。

以上のことは、すなわち、バッファ間の転送だけを考えれば蓄積しているデータの時間分以下の時間しか要しないが、最初の入力と、最後の出力がある限り途中のバッファでデータが待たされることになる。

なお、図では、各々のバッファの個数が2つの場合について書いているので、式(1)で示した遅延時間がミキシング回数ごとに蓄積されていく。この場合、 $\mu = 110$ msec、 $T = 100$ msec であるから、 $\mu - T = 10$ msec ごとに蓄積され、やがて、式(3)'で示す n 回目のミキシング処理後、すなわちここでは、 $n = 10$ 回目終了ごとに受信バッファからの出力が間に合わなくなり、強制的にバッファ内容が廃棄される。図中では、Pa13で示す音データが到着した時点でこの条件に合致する。これは、音声品質という、音の途切れ(skip)の現象となって現れる。ただし、現象が現れるのは、対象となる音データが鳴るべきときであるので、もう1回ミキシングが終了した後となる。

3.2 出力過剰モデルにおけるシミュレーション

このモデルでは、バッファサイズ T よりミキシング周期 μ が短いため、ミキシング対象となる音データが届いていない(生成はされているが伝送するバッファサイズ分溜まっていない)ことになり、このときにミキシング対象から外れてしまう。しかし、発生源の音は連続して発生しているので、1回のミキシング周期 μ 時間分だけ音の空隙(aperture)を生じる。遅延はこのとき1バッファサイズ T 時間分増える。また、送信端末ごとに空隙を生じるタイミングは異なる。

一方、出力側では、現在出力されている音が鳴り終わる前に次のミキシング周期が到来し、出力するための音が準備されるので、つねに供給過剰の状態になっている。したがって、現在出力中(鳴っている)音が出力され終わる前に、次のまたその次のミキシング終了音が準備されたとき、バッファに蓄積できなくなって、音データが廃棄される。このとき音の途切れ(skip)が生じる。

以下、数式で示す。ミキシングのタイミングは $T - \mu$ の時間ずつ早まる。 n 回ミキシングして、

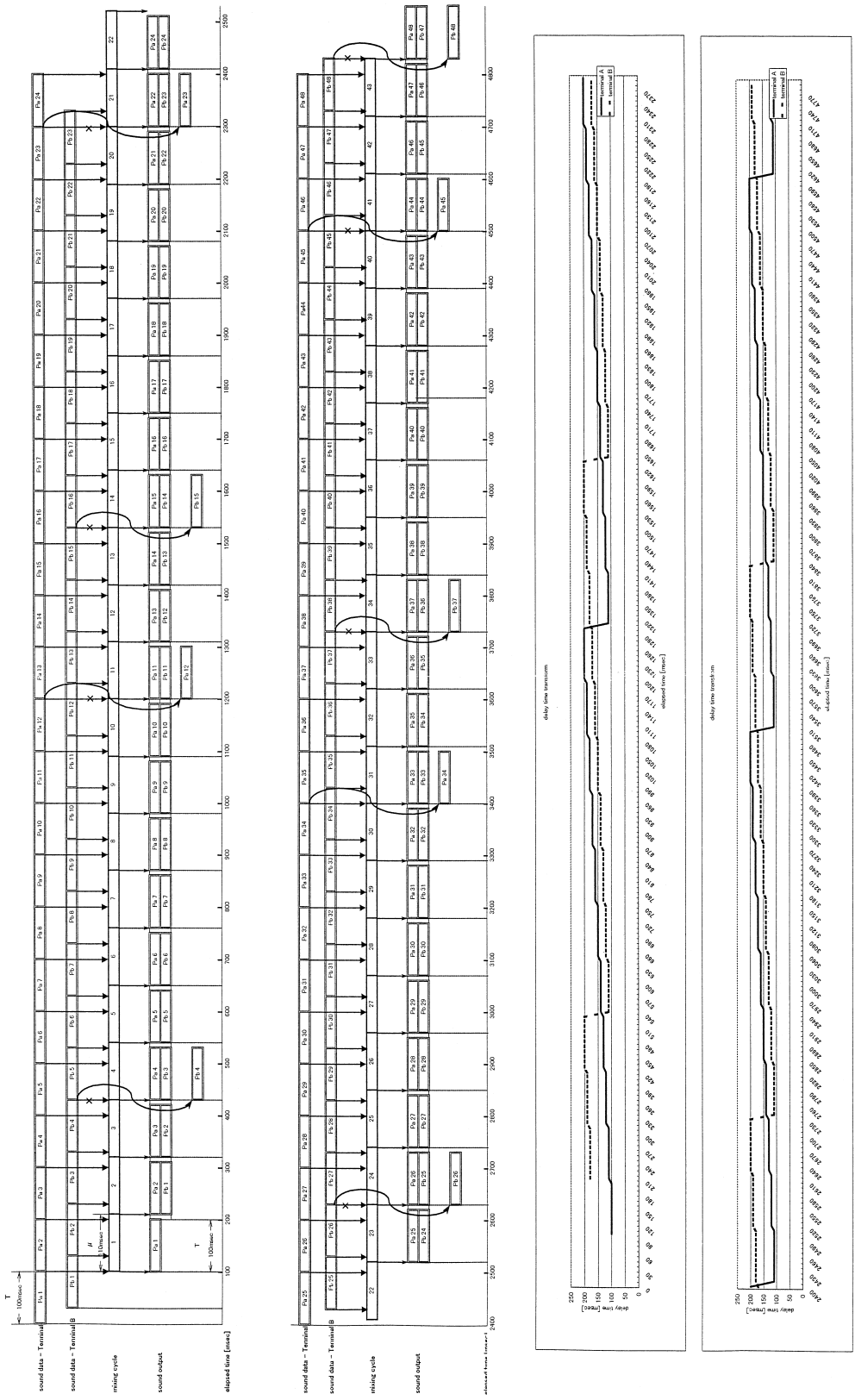


図 4 枯渇モデルにおけるシミュレーション
Fig. 4 Simulated change of delay time on "exhausted" model.

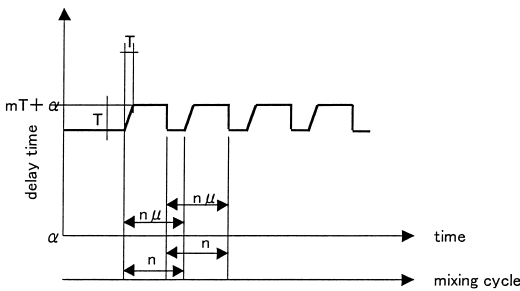


図 5 出力過剰モデルにおける遅延時間の推移

Fig. 5 Change of delay time on "surplus" model.

$$n(T - \mu) \geq T \quad (8)$$

となったとき、すなわち、

$$n \geq T/(T - \mu) \quad (9)$$

のとき、ミキシングのタイミングが来ているのに音データが到着していないことになり、音の空隙が生じる。すなわち、

$$n\mu = \mu T/(T - \mu) \quad (10)$$

の時間間隔ごとに音の空隙が生じる。しかし、この間、音は連続して出力されているので、遅延は増加しない。

また、まったく同じ $(T - \mu)$ ずつ、出力音が過剰になっていき、 $n\mu$ 時間ごとに出力しきれなくなり廃棄が生じる。このとき音の途切れ (skip) として現れる。遅延はこのとき、一挙にバッファサイズ T 時間分だけ短くなる。つまり、空隙が生じて遅延がステップ状に T 増え、途切れが生じてステップ状に T 減ることを $n\mu$ の周期で繰り返す。

音の空隙と音の途切れの周期は同じであるが、それぞれのタイミングは異なり、また、送信端末によっても異なる。

枯渇モデルと同様に、ここまではバッファが 2 つの場合について述べたが、 m 個あるときに拡張すると、以下ようになる。すなわち、式 (9) は、

$$n \geq mT/(T - \mu) \quad (11)$$

となる。これを模式的な図で示すと、図 5 のようになる。

過剰モデルのシミュレーションを図 6 に示す。ここでは、バッファサイズ $T = 100$ msec、ミキシング周期 $\mu = 90$ msec として示している。その他の条件は枯渇モデルの時と同様である。

式 (9) で示した、ミキシング周期 n 回ごと、ここでは、 $n = 100/(100 - 90) = 10$ 回ごとに、それぞれ別々のタイミングで、音の空隙と途切れが生じている。また、ミキシングのタイミングとバッファ受信契機が 2 台の送信端末で異なるので、それぞれの送信端末からの音で空隙および途切れの発生契機が異なっている

ことが分かる。図では、送信端末 A では、空隙が生じた後 4 サイクル目が終了した時点で途切れが生じているが、送信端末 B では、7 サイクル目で生じている。また、両端末について空隙が生じるまでは 1 バッファ分同期がずれていることが分かる。これはバッファサイズ T が大きいとコミュニケーションに対する影響が大きいと考えられる。

以上、枯渇モデル、過剰モデルについてシミュレーションを行ったが、 $\mu = T$ とするクリティカルな条件も当然想定できる。この条件を採用すると、理論的には、途切れも空隙も生じない。しかし、本論文では考慮していないが、伝送路の遅延などの外乱が生じたときに、対処することが困難になると思われる。到着をトリガとしてミキシング処理を行うという方法もとれるが、外乱に応じて音の空隙が生じるということになる。

4. 実験

シミュレーション予測が正しいかどうかを判定するために、実測実験を行った⁵⁾。実験環境を図 7 に示す。送信端末は 2 台用意し、受信端末は 1 台とした。これらを 10 Mbps の Ethernet LAN を用いて接続した。LAN 上にはこれら以外の端末は接続せず、伝送路に品質の影響を与えないようにした。これ以外に、測定用の端末を 1 台用意し、これに、1 台の送信端末への入力音をバイパスした音と、受信端末の出力音を入力してモニタした。バイパス音と受信端末の出力音は共にアナログ音である。両者の間のずれを波形解析ツールを用いて見ることで遅延時間を測定した。送信端末から送られる音は、パワー (音量) の異なるパルス音で、これをそれぞれ右チャンネル、左チャンネルに入れて伝送することで、測定をやりやすいようにした。使用した端末はすべて、Pentium 200 MHz、メインメモリ 32 MB のパソコンである。

送受信バッファともサイズ T は 1 つあたり 120 msec で、個数 m は 10 個とした。音データはすべて、8 kHz サンプリング、8 bit 量子化を行っている。これを 10 Mbps の Ethernet で伝送する。

図 8 に枯渇モデルの条件で行った実験の実測値のグラフを示す。ここでは、ミキシング周期 $\mu = 132$ msec である。また、図 9 に過剰モデルの条件で行った実験の実測値のグラフを示す。ここでは、ミキシング周期 $\mu = 108$ msec である。これら図中の delay-1 ~ delay-5 は各々 5 つのサンプル音においての計測値である。

各々のグラフは、測定した時点が間欠的であるため、図 3 や図 5 に示したモデル図や図 4 や図 6 に示した予測シミュレーション図そのものにはなっていないが、それ

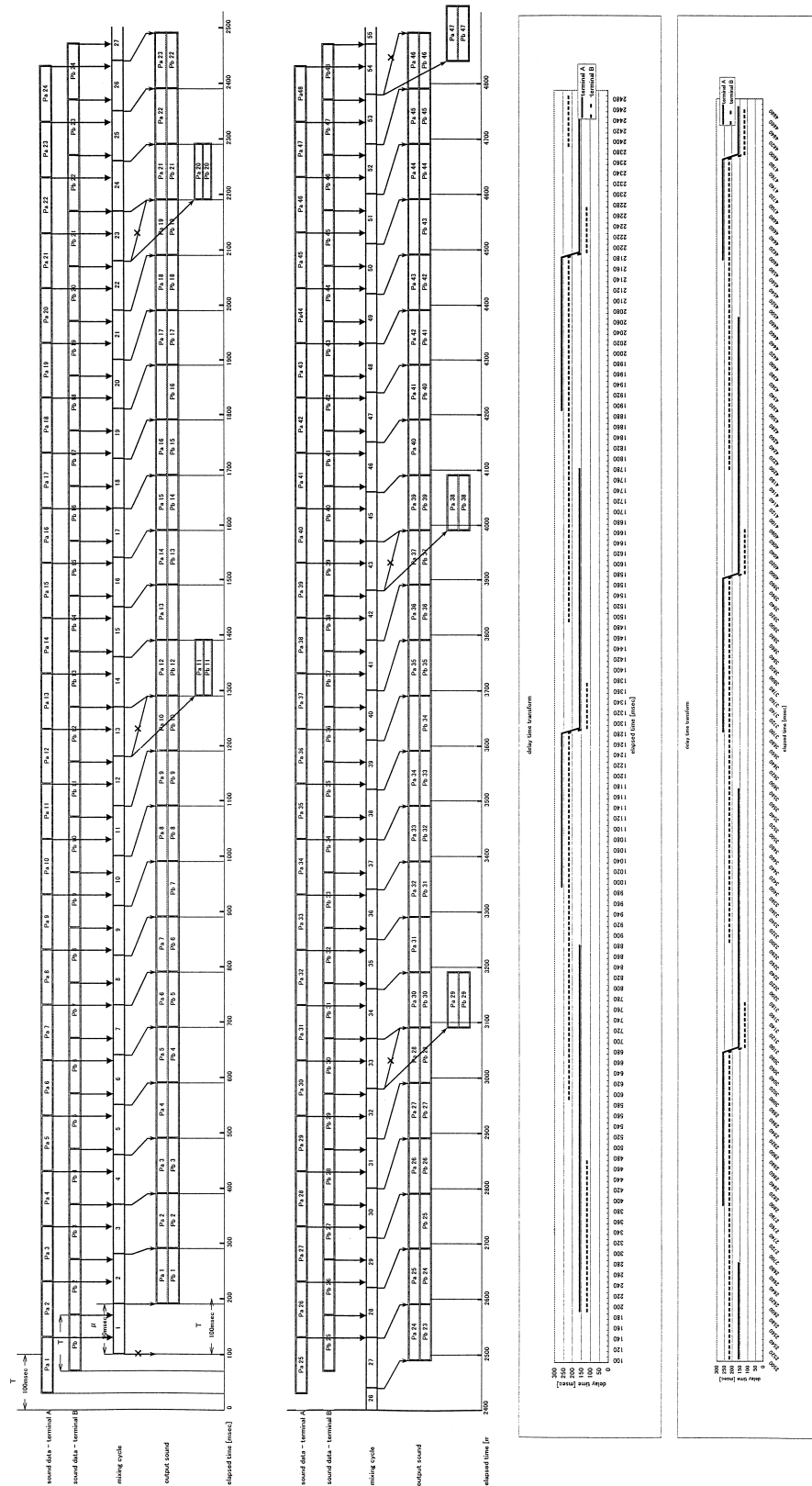


図6 過剰モデルにおけるシミュレーション
Fig. 6 Simulated change of delay time on "surplus" model.

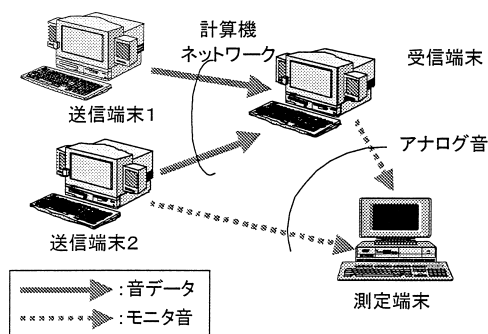


図 7 実験環境

Fig. 7 Experiment configuration.

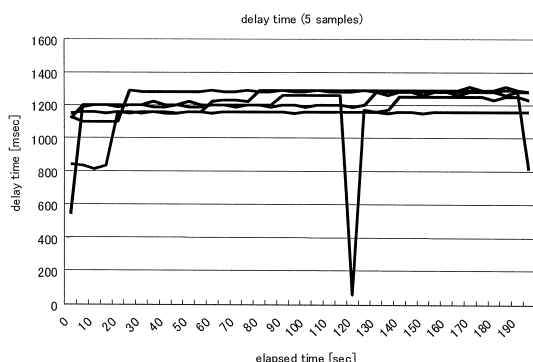


図 8 出力枯渇モデルにおける遅延時間の実測値

Fig. 8 Experimented change of delay time on "exhausted" model.

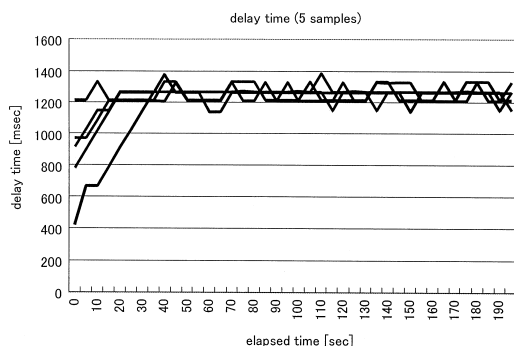


図 9 出力過剰モデルにおける遅延時間の実測値

Fig. 9 Experimented change of delay time on "surplus" mode.

それぞれのグラフをサンプリングしたときの図であることが分かる。すなわち、枯渇モデルにおいては、図 3 における最大遅延時間 $mT + \alpha$ と極小遅延時間 $(m-1)T + \alpha$ の差 $(mT + \alpha) - \{(m-1)T + \alpha\} = T = 120 \text{ msec}$ であり、同図では所々に送信音不連続によると思われる急激なゼロクリアはあるものの、安定周期部分のサンプル区間における幅となっていることが分かる。各

サンプル内での周期、すなわち、式 (5) から得られる 12 sec はグラフからは残念ながら見出せない。また、過剰モデルにおいては、図 5 における振幅 $T = 120 \text{ msec}$ および周期 $n\mu = m\mu T / (T - \mu) = 10.8 \text{ sec}$ であることが図 9 から見る事ができる。

5. 考 察

本章では、バッファとミキシングの適切な設定について考察する。

(1) バッファの個数: m

バッファの溢れ対策を講じるなら、受信側のバッファの個数(サイズではなく)を増やすことが考えられる。しかし、バッファの個数を増やすと遅延は累積していく。すなわち、受信側で聴く音が送信側では過去の音となるため、コミュニケーションへ与える影響は大きい。入出力のボトルネックに陥らない程度の最小限の個数にすべきである。

(2) バッファのサイズ: T

バッファ 1 つのサイズは、小さい方がよい。いずれのモデルにおいても、遅延の変動分に相当するためで、これが大きいと変動も大きくなる。

(3) ミキシング周期: μ

ミキシング周期は短い方がよい。なぜなら、シミュレーション時に α で示した、送信端末ごとのオフセットの差、すなわち遅延時間の違いが大きくなることは避けた方がよいからである。

以上、すべての条件を満たすなら、バッファサイズ T は小さく、ミキシング周期 μ も短く、バッファの個数 m も少なく、ということになる。しかし、 T も μ も小さくすると、処理系への負荷が大きくなり、オーバーヘッド分が無視できなくなるとされる。

(4) バッファサイズとミキシング周期の関係: T, μ

枯渇モデルと過剰モデルのどちらが良いかということであるが、ここでは、相互にやりとりを行う会話に与える影響の観点から考察してみることにする。

過剰モデルではバッファの廃棄が生じて音の途切れ (skip) が生じるまでは、送信側の音の同期がずれている状態(送信端末ごとの遅延時間が異なっている状態)が生じる。これもコミュニケーションに与える影響が大きい。また、過剰モデルでは、不均衡に音の途切れと空隙が生じるので、これが送信端末ごとに異なる結果をもたらす場合がある。実測実験時には生じなかったが、図 10 に示すような、送信端末 B にとっての途切れ時に送信端末 A の空隙が生じるようなタイミングの場合、送信端末 A では一定の遅延(バッファサイズ分)しかないにもかかわらず、送信端末 B では

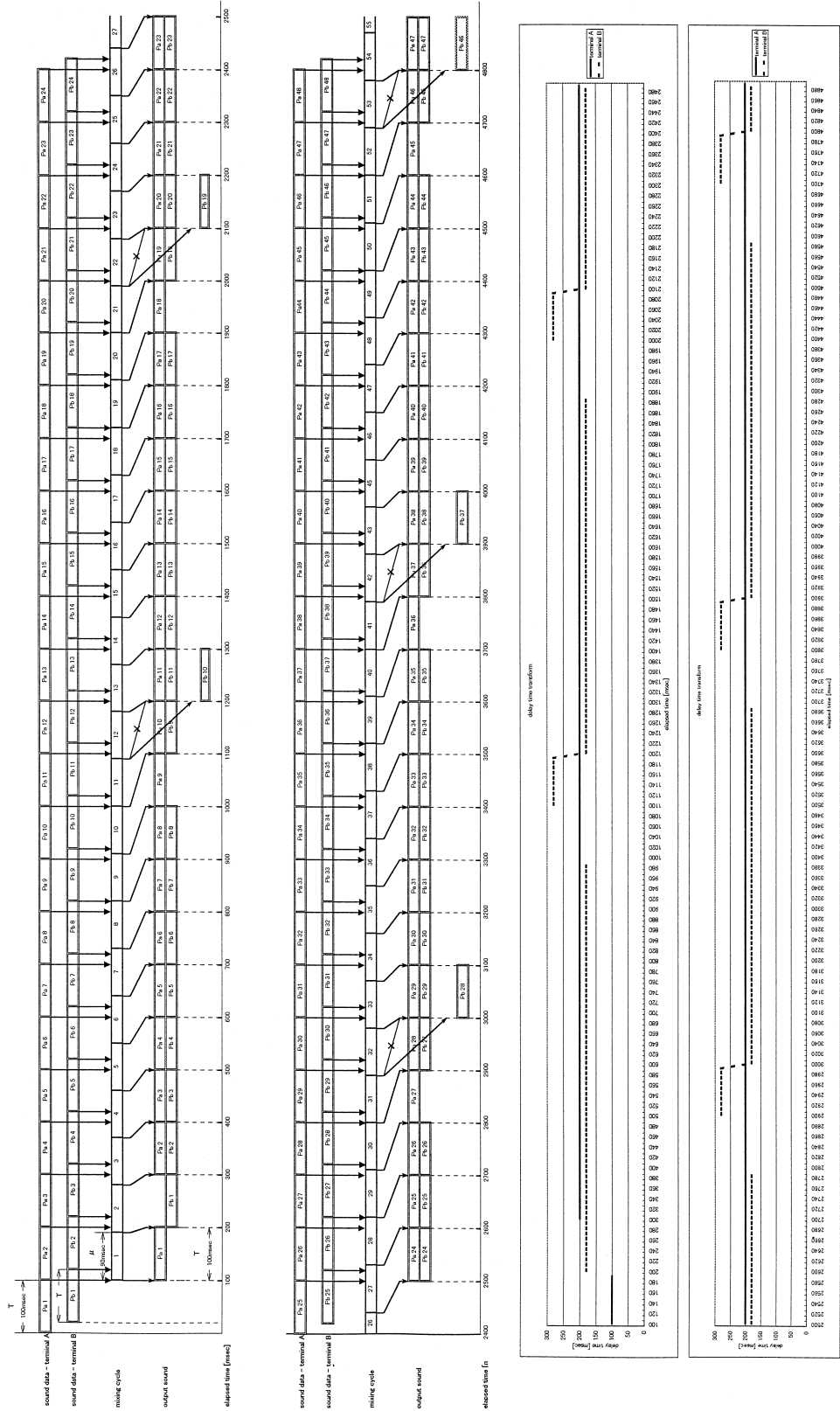


図 10 過剰モデルにおける特異な現象例 (シミュレーション)
Fig. 10 Extraordinary case on "surplus" model (simulated).

音の途切れと空隙が一定周期で現れるという不均衡な状態になる。

逆に、音の品質を下げってしまうことになるが、その低下の度合いが平均的である、枯渇モデルが人間にとって予測しやすいパターンなので順応しやすいともいえる。バッファサイズ T とミキシング周期 μ を小さくした枯渇モデルの採用に期待が持てる。

以上、音のミキシングが存在するときの設定条件について考察してきたが、ミキシングの必要性は、多人数参加型の状況がビジュアルに示されているときに重要になってくるものである。チャット形式でビジュアルな呈示が存在しない、多人数参加型の音声会話が行えるサイバースペースでは、ミキシングをせずに順次音声出力させるだけにしている。ある程度のリアル性を排除しても人間が慣れてしまうわけであるが、これは、マルチメディアコミュニケーション環境において、特定のメディアのリアル性を追究すると、他のメディアもこれに付随してリアル性を追及せざるをえないということを示唆しているといえる。

6. ま と め

複数端末からデジタル音を連続的に送信し、受信端末側でミキシング出力するシステムを考えたときに、ミキシングの周期と受信側のバッファのサイズ（蓄えられる音の長さ）の関係で、音の途切れと空隙が生じることをシミュレーションと実測実験で示した。ミキシング周期がバッファサイズより長い場合（出力枯渇モデル）では、定常的に音の空隙を生じるとともに、定期的に音の途切れを生じ、ミキシング周期がバッファサイズより短い場合（出力過剰モデル）では、異なるタイミングではあるが同じ周期で定期的に音の途切れと空隙が生じることを示した。また、これらの現象が不可避であることをふまえて、より適切な設計条件について考察を行った。複数端末からの送信かつ連続的な送信という制約の下ではあるが、より多くの送信端末が同時に存在するアプリケーションを想定した場合、近似的な条件となる。

今回は、受信端末 1 台だけの現象に注目したが、今後は、相互に送受信を行っている個々の端末間の違い（絶対時間に対する個々の端末間の遅延のずれや音品質の違い）についてシミュレーションを行い、適切な設計条件を検討する。また、この設計条件を反映した多人数参加型コミュニケーションシステムの構築を行って、主観評価を行う。

謝辞 執筆にあたり励ましてくれた NTT インテリ

ジェントテクノロジー正木茂樹部長に感謝いたします。また、開発と実験およびディスカッションに協力してくれたインタースペース開発関係の皆様にも感謝いたします。

参 考 文 献

- 1) Hayashi, Y. and Koyano, H.: Technological Issues on VoIP and new directions of applications, *IBC Voice over IP Internet Telephony '98 Conference* (May 1998).
- 2) 星, 谷川, 松井, 岩見, 寺田: LAN 環境における負荷適応制御を用いた低遅延リアルタイム音声通信システム, 情報処理学会論文誌, Vol.40, No.7, pp.3063-3073 (1999).
- 3) 湯田, 清末: サイバースペース環境を用いた仮想オフィスシステムの提案, 情報処理学会研究資料, 97-GW-21, No.13, pp.145-150 (1997).
- 4) 安藤, 阪谷, 渡辺, 林, 小谷野: インターネット電話 Vocalink のアーキテクチャ, *NTT R&D*, Vol.46, No.9, pp.977-984 (Sep. 1997).
- 5) 湯田, 森内, 清末, 正木: 仮想オフィスシステムにおける音声通信方式に関する検討, 第 58 回情報処理学会全国大会論文集, 1Q-5, Vol.3, pp.441-442 (1999).

(平成 12 年 2 月 25 日受付)

(平成 12 年 9 月 7 日採録)



清末 悌之 (正会員)

昭和 35 年生。昭和 58 年九州大学工学部電気工学科卒業。同年電電公社入社。ヒューマンインタフェース研究所画像情報メディア研究部、企業通信本部システムインテグレーション部等を経て、現在 NTT 東日本通信機器事業部第一商品部第 2 プロジェクトマネージャ。文書画像処理、動画像ハンドリング、3 次元サイバースペース内でコミュニケーション場の提供等の研究に従事。電子情報通信学会、日本バーチャルリアリティ学会各会員。



湯田 佳文 (正会員)

昭和 40 年生。昭和 63 年鹿児島大学工学部機械工学第二学科卒業。同年日本電信電話 (株) 入社。通信設備の保守・設計、仮想空間における多人数間コミュニケーションについての研究会開発に従事。現在西日本電信電話 (株) 勤務。