

デスクトップ会議システム Open GroupMedia における音声会話*

4E-4

澤島 信介, 伊藤 清司, 長谷部 浩一, 村永 哲郎†

(株)東芝 研究開発センター 情報・通信システム研究所 第二研究所‡

1 はじめに

我々は既存の個人用アプリケーションを同時に共用し、かつ音声などの各種メディアを支援するグループウェア環境 Open GroupMedia (以下 OGM と書く) の研究開発を行なっている [1]。OGM では共有ウィンドウ機能がすでに動作しているが、今回さらに多者間音声会話機能を付加した。本音声会話システムの特徴的機能と、無音部削除による音声の遅延解消についての評価結果を報告する。

2 音声会話システム

2.1 特徴

本音声会話システムは、次の特徴を持つ。

- (1) 電話に似た使用感を持ち、多者間会話ができる
- (2) OGM における共有ウィンドウ部と独立したシステムであり、音声会議参加者情報を共有ウィンドウ部へ転送する機能を持つ

従来から、一般の音声会話手段として電話が利用されている。そしてワークステーション上で多者間の音声会話機能を実現する場合も、利用者に馴染みのある電話に似た使用感を与えるインタフェースを持つことが有効と考えられる。よって本システムでは電話に似た使用感を持たせることを設計方針とした。

次に、わざわざ共有ウィンドウの機能を使わなくても、音声だけの会議で事が足りる場合が多い。そこで、共有ウィンドウ機能と独立にすることで音声だけの会議を可能とした。しかし、そこから共有ウィンドウ会議へ移行する必要がある場合もあり、それを容易にするために音声会議参加者情報を共有ウィンドウ部へ転送する機能を設けてある。

2.2 多者間会話

本システムでは、電話の機能に加えて次の機能を提供することで多者間会話を可能とした。

- (1) 1人の利用者は複数のセッションに参加することができる。ただし、複数のセッションに参加している場合はどれか1つのセッションでのみ会話できる。
- (2) 割り込み通話により新たなセッションが結べる。

*Audio Communication for Desktop Conference System Open GroupMedia

†N.Sawashima, K.Ito, K.Hasebe, T.Muranaga

‡Communication & Information Systems Research Laboratories, Research and Development Center, TOSHIBA Corporation

- (3) 2つのセッションを融合して1つのセッションとすることができる。

ここで、セッションとは複数の利用者からなる会議の単位のことである。割り込み通話とは、すでにいくつかのセッションに参加している状態で、さらに他者からの呼び出しを受けるか他者を呼び出すことをいう。

多者間会話は (2) と (3) を繰り返すことで可能となる。

2.3 構成

本システムは LAN に接続された UNIX ワークステーション上に構築された。

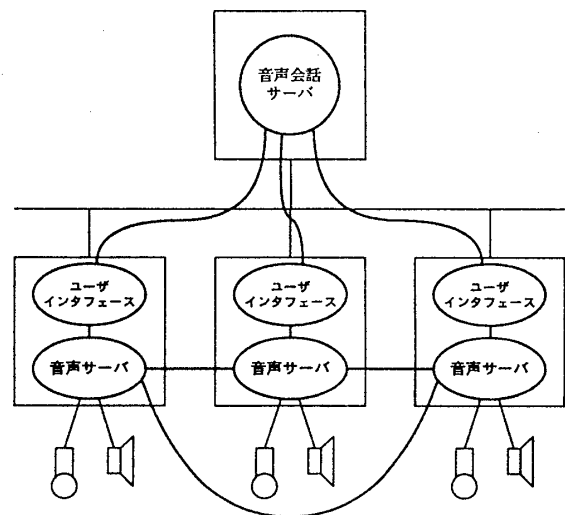


図1: 音声会話システムの構成

音声入出力デバイス: 入力された音声は 8Kbyte/sec モノラルの品質に A/D 変換される。出力はその反対となる。入力用、出力用にそれぞれ 1 秒 (8Kbyte) のバッファを持つ。

音声サーバ: 各ホストマシンにある音声入出力用デバイスを管理する。また、音声サーバ同志で互いに音声データを送受信し、ミキシングする機能を持つ [2]。内部に合計約 1Mbyte の音声データ蓄積用バッファを持つ。音声の遅延を解消するために無音部を検出し削除する機能 [3] も組み込まれている。

ユーザインタフェース: ユーザインタフェースとしての機能の他に、音声通信の接続・切断、音声転送の

開始・停止などの制御命令を音声サーバに送る機能も持つ。

音声会話サーバ: セッションの管理をし、音声会話機能の実現をする。

3 遅延時間に関する評価

3.1 目的

音声会話システムにおいて、音声入力してから応答を得るまでの遅延は使い心地に影響する重要な要素である。一般に UNIX マシンでは様々なプロセスがバックグラウンドで走っており、それらによる負荷状態と音声転送の遅延時間の関係を調査することにした。

同時に、本システムが遅延解消のために導入した無音部削除機能の有効性について検証する。

3.2 方法

イーサネットに接続された2台のワークステーション(東芝製 AS4075: 21.8SPECint)において、一方に音声を入力し、他方からその音声を出し、入力音声と出力音声を同時にステレオ録音する。その時、無限ループのプロセスをバックグラウンドに複数個走らせ、その個数を調節することにより CPU 負荷を制御する。録音終了後、その内容から経過時間に対する遅延時間を測定する。測定方法としては、簡単のため入力音声と出力音声をグラフ化し、ずれを目視することにより行う。

以上の操作を無音部削除機能を動作させた場合とさせない場合について行ない、比較する。

なお、負荷状態を数値化するために負荷平均を使用した。負荷平均とは UNIX で使用される CPU 負荷の指標となる値で、プロセスキューに存在するプロセス数の平均値である。

3.3 結果と考察

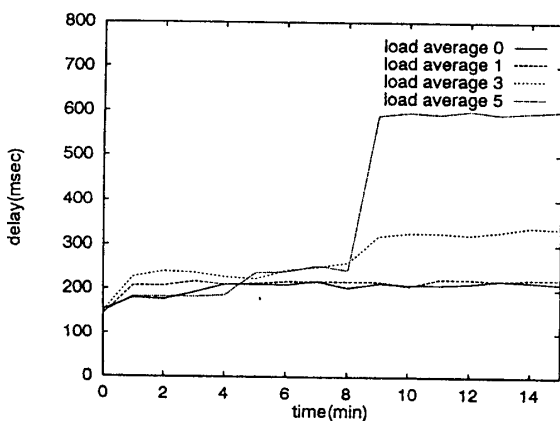


図 2: 無音部削除をしない時の遅延時間

図 2 より、次のことがわかる。

- 最小遅延時間は約 150msec。ここで、最小遅延時間とは音声データが入出力デバイスやネットワー

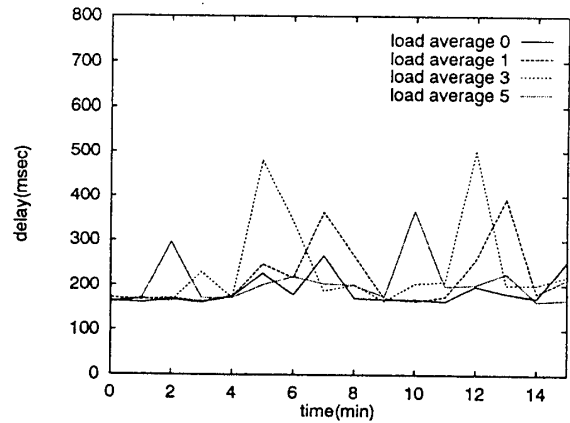


図 3: 無音部削除をした時の遅延時間

クなどを通過するために必要な最小の時間のことである。

- 遅延時間は瞬間的に蓄積する。本システムの実行環境におけるプロセッサの処理能力やネットワークの転送速度を考慮すると、音声転送レート 8Kbyte/sec は十分処理可能なはずである。遅延が瞬間的に蓄積する理由としては、

- デモン・プロセスがスケジューリングされる
- ネットワーク・トラフィックの増加

が考えられる。

また、図 3 では無音部削除機能が有効に働き、蓄積された遅延が解消することがわかる。このときの音声出力を聞いたところ不自然な音の切れなどによる違和感はなく、音質の面からも有効と感じられた。

4 おわりに

OGM における、電話に似た使用感を持つ多者間音声会話システムについて報告を行なった。また、遅延時間を評価した結果、本システムで採用している無音部削除による遅延解消が有効に動作していることが確認できた。今後は本音声会話機能をより進化させるとともに、OGM に動画機能を付加するなど、さらに使いやすいグループウェア環境とすべく模索していきたい。

参考文献

- [1] 村永 哲郎, 他. Open GroupMedia — 個人用アプリケーションを共有するグループウェア・フレームワーク, 情報処理学会グループウェア研究会, Vol.GW-4, No.18, pp.121-128, October 1993.
- [2] 長谷部 浩一, 他. UNIX ネットワークにおけるオーディオサーバ, 情報処理学会第 44 回全国大会, 3, pp.323-324, 1992.
- [3] S. Wing Keung To, 他. A Voice Detection Method for Power Saving System in Digital Cellular Phone, 1993 年電子情報通信学会春季大会 A-224.