

仮想空間における多人数会話制御の検討

5M-9

A Study of Audio Control for Multi-User Virtual Space

*森内 万知夫

*小長井 俊介

**大井 充

*NTT ヒューマンインタフェース研究所

**NTTソフトウェア（株）

1. はじめに

インタースペース（以下、IS）は、3次元の仮想空間を複数の端末利用者がネットワークを介して共有し、映像・音声・テキストをベースにリアルタイムコミュニケーション環境を構築するプラットフォームである。なかでも、音声による会話は、毎日、インターネットの利用者の間で活発に行われている。しかし、この音声会話では、従来1つの仮想空間内では同時会話人数が2人迄という制限があり、2人の会話中に3人目以降の利用者が途中参加することが難しかったり、2、3人での閉じた会話形態になりがちであった。そこで、本稿ではより汎用的な会話形態に対応するため、1つの空間内で3人～5人の利用者が同時に会話できる低ビットレートでの音声システムを試作したので報告する。

2. 音声制御法

仮想空間の世界で、コミュニケーションを行うには、音声、映像、テキスト文字、分身のジェスチャー等の手段があるが、現実の世界と同様に会話をリアルタイムに行うには、音声通信とその制御が重要になってくる。

空間上での音声制御は、ISでは現在、同じ仮想空間にいる利用者には、その位置に無関係に2人の利用者の音声をミキシングして、発言者以外にサーバーから送信するという、平面的な音声制御である。

従来の利用者2人だけの同時ミキシングだけでも、10～20人の利用者が集う井戸端会議

的な会話の場合では、会話が途切れた段階で会話が切り替わるよう、自然にフロア制御されている。この点に関しては、一般に電話での1対1での通話になれていること、聞き手も発話者の数に限界があること等が要因の一つと考えられる。

しかし、現状のISでは、不特定多数での会話を行うので、同時に3人以上の利用者が偶然発言したり、3人目以上の利用者が一斉に笑い声等を発声することが起こる。

このような場合には、3人目以降の声は削除され、発言した内容が他者に伝わらなかったり、同じ空間にいるという場の同一性や場の雰囲気等が伝わらない。

また、従来の音声のミキシング方式では、CPUの負荷率が高く、遅延や音切れの要因となっていた。本方式では、これらの問題点について、低負荷で低ビットレートでの音声通信方式を検討した。

3. システム概要

システム構成図を図-1に示す。音声サーバーに複数のクライアントがLANまたは、回線により接続されている。サーバーは、各クライアントからの音声をデコード/ミキシング/エンコード処理した後、全クライアントに配信する。これは、クライアントの負荷をなるべく軽減するためである。

サーバーは、サンプリング周波数等の音声コーデックに必要なパラメータを一括管理する。クライアントの無音レベルの調整等の通信パラ

メータは、サーバー接続時に各クライアントの音声入力装置のレベルを測定して決定されるので、クライアントでは音声の入出力量のみを調節する。(図-2)

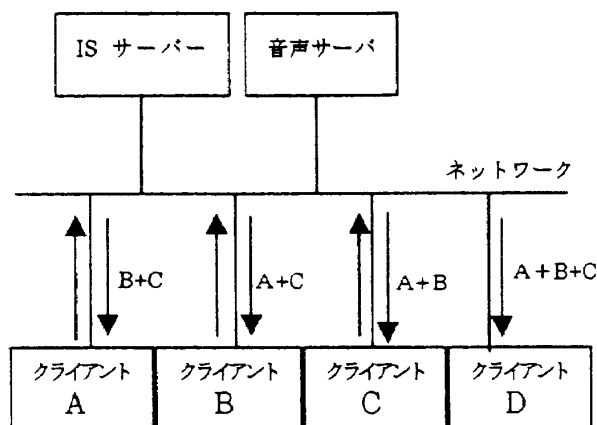


図-1 システム構成と音声の配信



図-2 クライアントの制御表示

4. 音声通信方式

音声コーデックは、CPUの演算量削減のため、ADPCM方式で、サンプリング周波数5kHzで行っている。

音声サンプル値は、増大や縮小といった変化が継続的続く傾向にあり、この変化の間は同じ符号が続くことになる。この点を考慮すれば、量子化処理では同符号が続く場合は、この符号は処理せず、音声サンプル値の差分量が変化して符号が変化する場合のみ処理できれば、情報量を削減できる。本方式では、量子化ビット中に符号制御コードを1ステップもうけて、音声サンプル値の差分の変化時にこのコードをエンコード時に量子化データに挿入し、デコーダ側でこの制御符号を基に差分量を計算した。これに

より、従来の量子化精度を維持しつつ、量子化1ビット分を削減することができた。

サーバーのミキシングの処理は、ある規定の時間内に音声サーバーで受信した音声パケットを同時発声とみなして行う。但し、規定時間内でも同じクライアントからの音声パケットを受信した場合、このパケット以前を同時発声とみなす。また、自分の音声はエコーバックしない。つまり、サーバーはA, B, Cが同時発声した場合、AにはB+C、BにはA+C、CにはA+B、その他の利用者にはA+B+Cの4種類のミキシング処理を行う。

4. インタースペースでの実装

従来方式と本方式との伝送速度、音声入力からエンコード・デコード後、音声出力するまでのCPUの使用率と遅延時間の比較を行った。^[3] その結果、現行方式と比べ、CPU使用率、伝送速度、遅延時間ともに改善し、多人数でのミキシング実現の見通しを得た。ISへの実装では、LAN上で5台のクライアントを接続して会話実験を行い、音質は現行方式にやや劣るが、遅延時間の差はなく会話ができた。

5. むすび

本方式により、LAN上での動作では、最大5人での同時ミキシングが可能となった。また、低ビットレートでの音声伝送とCPU使用率の削減を実現した。今後は、音声品質の向上について検討したい。

6. 参考文献

- [1] 谷川他：“多人数参加型仮想環境における音声制御法”，学技法，IE95-43，1995-07
- [2] S. Sugawara：“InterSpace”，IEICE Trans.，Vol. E-77-D，No. 12，1994
- [3] 大井他：“インタースペース音声通信方式の検討”，信学大全 B-8-90，Mar. 1998