

協調 VR 環境のための音声通話ライブラリ

久木元伸如 †¶ 江原 康生 ‡ 小山田 耕二 §

東和大学情報学科 † 〒 815-8510 福岡市南区筑紫丘 1-1-1

京都大学大学院工学研究科¶ 学術情報メディアセンター ‡ 〒 606-8501 京都市左京区吉田本町

京都大学高等教育研究開発推進センター § 〒 606-8501 京都市左京区吉田二本松町

要旨

近年、遠隔の共同研究者との協調作業の重要性は高まっており、可視化分野においても VR(Virtual Reality) 空間を共有した協調作業が注目されている。本稿では、広域ネットワーク網における協調 VR 環境の構築に向けて、VR 共有空間における高品質な音声通話によって、円滑なコミュニケーションを実現するために開発した音声通話ライブラリについて述べる。本ライブラリは VR 空間における発話者との位置関係に応じて動的に音量と PAN を調整することが可能である。さらに HybridP2P 型のネットワークを構築し、通信路毎に異なるポートを割り当てることによって、MCU や IP マルチキャストを用いずに多地点同時通話を実現することができる。学内 LAN での実験において、多少の遅延は発生するが、高品質な音声通話が可能であることを示した。

Voice Communication Library for Collaborative Virtual Environment

Nobuyuki KUKIMOTO†¶ Yasuo EBARA‡ Koji KOYAMADA§

Department of Information, Tohwa University† 1-1-1 Chikushigaoka, Minami-ku, Fukuoka 815-8510

Faculty of Engineering¶ Academic Center for Computing and Media Studies‡, Kyoto University

Yoshida-honsho, Sakyo-ku, Kyoto 606-8501

Center for the Promotion of Excellence in Higher Education§, Kyoto University

Yoshida-nihonmatsucho, Sakyo-ku, Kyoto 606-8501

Abstract

In recent year, the construction of collaborative works between remote places has been recognized important. Moreover, the demand for the collaboration by sharing the virtual environment in a field of data visualization has been rising. In this paper, we developed an audio library for voice communication (vocAL) to realize smoothy communication in a collaborative virtual environment over the wide area network (WAN). The vocAL can dynamically adjust the sound volume and the PAN level correspond to the position of a speaker in VR environment. To realize a voice communication in multipoints without a MCU and IP multicast, we constructed a HybridP2P network and proposed a method which assigned different port number for each communication path. In this experiment via campus LAN, we showed high-quality voice communication, while a problem of delay generated.

1 はじめに

近年、科学技術分野では計算機インフラとして、地理的・組織的に分散したコンピュータやデータベースなどを統合・接続した環境が整備されつつある。このような環境を利用し、各研究機関が協調して研究開発に取り組む動きが広まる傾向にある。特に大規模な可視化データを扱う研究分野では、遠隔地間で可視化データを参照しながら、共通の問題に取り組むことが非常に有用とされる。それに伴い、遠隔の共同研究者やユーザが Virtual Reality (以下 VR) 空間を共有しながら協調作業を行う環境 (以下 CVE: Collaborative Virtual Environment) に関する研究が行われている [1],[2]。

我々は、新たな知識創造を支援する CVE の実現を目指し、広域ネットワーク網を通じて VR を用いた臨場感のある共有可視化空間の構築を目指している。協調作業を行う場合には、双方向によるコミュニケーションが不可欠であり、そのコミュニケーションの質が作業効率を大きく左右するものと考えられる。すなわち、的確な意思の疎通の実現が重要となってくる。一方、協調作業の際のコミュニケーションは一般的に会話の占める割合が大きい。そのため、VR による可視化空間でも円滑な協調作業を行う上で、音声通話などを利用した会話によるコミュニケーションが不可欠とされる。

本稿では、協調 VR 環境における円滑なコミュニケーションを支援するための音声通話ライブラリを開発し、VR 空間に実装を行った。その概要について述べる。

2 協調 VR 環境

VR は知的発見を行うための有用な可視化技術の 1 つである。しかし、没入型多面ディスプレイ (以下、IPT) による VR システムでは、可視化に必要なとされる操作や制御を行う際に、ホストコンピュータに戻って操作を行う手間が必要となる。それによって、作業に多大な支障を与えるなどの問題が生じる。一方、3次元画像によって VR 環境を構成する可視化空間に注釈付与が可能になれば、利用価値の高い協調環境が実現できると考える。我々は、PDA を用いた VR 用の入力ユーザインタフェースを開発し、VR 環境に実用レベルでの実装を行った [7]。

実世界で複数の人間が一つの問題に取り組む際

に、各自の経験や知識を持ち寄り、様々な角度からの解析によって、新たな知識発見や解決方法の導出が可能になる。つまり、VR による可視化空間においても、複数人による協調作業を行う方が効率的であり、かつ有益な成果を得られやすいことが多い。しかし、現在利用されている VR による可視化システムは、個人による作業環境のサポートしか行われておらず、複数人での協調作業が困難である [3],[4]。

遠隔協調作業の事例として、ビデオ会議システムや AccessGrid[5] などがある。現在、インターネットを利用したビデオ会議システムは H.323 という ITU-T 勧告 [6] に準拠している。H.323 とは、呼制御や端末間のネゴシエーションなどを定めたプロトコル群で、例えば、Microsoft 社の NetMeeting、Polycom 社の View-Station や ViaVideo など多くのシステムが準拠しており、これらの相互運用が可能である。H.323 は基本的に一对一のビデオ会議のための規格であり、複数の拠点を同時接続することができない。複数の拠点で会議を行うためには MCU (多地点接続装置) が必要となり、アクセス可能なネットワーク上に MCU を設置することで多地点接続を実現している。一方、Access Grid は IP マルチキャストを利用しており、接続拠点数が増えてもメディアの品質が損なわれない利点がある。

しかし、MCU を用いた場合は拠点数が増えると音声などメディアの品質劣化が生じ、円滑なコミュニケーションが困難になる。また Access Grid は、IP マルチキャストを行うため、各拠点の接続ルータにマルチキャスト対応機器を設置する必要がある。そのため、機器を設置できない拠点は Access Grid に容易に接続できないので、協調作業に参加できなくなる問題が生じる。

本研究では、円滑なコミュニケーションを行えるように、高音質な音声通話ライブラリを開発し、協調 VR 環境に実装を行った。また、MCU や IP マルチキャストを用いずに多地点同時通話が可能な方式を提案する。次章以降、システム構成について述べる。

3 システム構成

3.1 VR 可視化システム

本研究では可視化アプリケーションの構築に VTK(Visualization Tool Kit) を利用し、グラフィッ

クスライブラリには OpenGL Performer を，VR ライブラリには CAVE ライブラリを利用している．システム構成を図 1 に示す．CAVE ライブラリはトラッキングデータとコントローラデータを管理する trackd デーモンを必要とする．trackd に入力されたデータは共有メモリに格納され，CAVE ライブラリは共有メモリを参照することによってトラッキングデータやコントロールデータを得る．本システムでは，可視化操作や可視化空間への注釈付与などのインタラクションを行うコントローラとしての操作用 PDA アプリケーションを開発し，PDA 上に実装している [7]．PDA 本体の正面 2 つのプログラマブルボタン，側面のプログラマブルボタンおよびカーソルキーからの入力をコントローラデータとして，通信機能の拡張を行った trackd に送信する．PDA には東芝製 Genio e 550GX (PocketPC 2002) を用い，無線 LAN アダプタを装着して，ホストアプリケーションおよび trackd との通信を行う．空間位置センサに Polhenus 社の Fastrak を用い，LongRanger で計測範囲を広くした．センサの位置や傾きを計測することによって PDA のトラッキングデータを RS-232C で trackd に送信する．

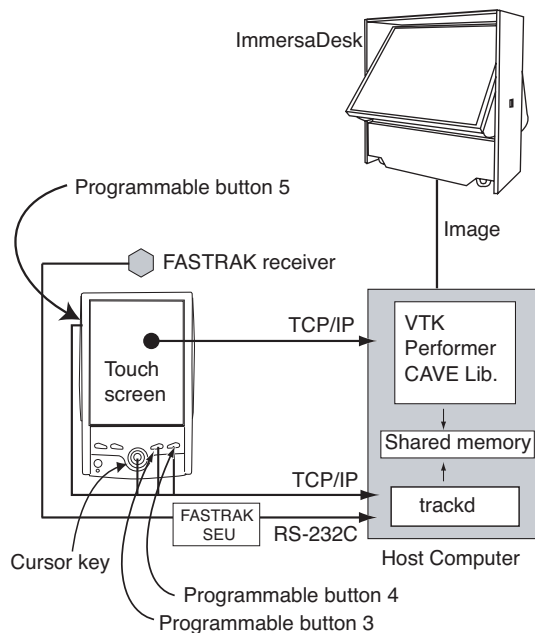


図 1: ローカルシステムの構成

3.2 Hybrid P2P

CVE では，ローカルと他のノードの VR 空間上の情報が同期して変化しなければ，整合性が崩れるため，空間共有が成り立たなくなる．この同期に関する問題を解決するために，ネットワーク上に管理サーバを設置して，VR 空間のイベントや状態を管理する server-client 型方式が考えられる．しかし，協調作業では必要に応じて遠隔地のユーザや専門家と動的なグループ形成ができることが重要とされる．server-client 型のネットワークでは，サーバが停止している際には，必要な時に相手と接続する動的なグループ形成ができなくなる問題が生じる．

よって，高いスケーラビリティを有し，動的なグループ形成が可能である P2P によるネットワークの応用が考えられる．本研究では，server-client 型と P2P 型の利点を兼ね備えた図 2 に示す Hybrid P2P 型 [8] のネットワークを用いる．

Hybrid P2P において，サーバは参加してきた協調者の IP アドレスを管理する．新しい協調者はサーバにアクセスし，自分の IP アドレスを登録し，既に共有ネットワークにアクセスしている協調者の IP アドレスをサーバから受信する．受信した IP アドレスを用いて，既に共有ネットワークにアクセスしている協調者へアクセスする．よって，各ノードはお互いの IP アドレスを事前に知る必要なく，共有ネットワークにアクセス可能となる．

3.3 音声通話ライブラリ

3.3.1 仕様

質の高いコミュニケーションを実現するためには明瞭な音声による通話が求められる．しかし，既存の協調 VR で用いられる音声通話は電話程度の品質しか保証されていない．しかし，近年のネットワークの広帯域に伴い，品質の高い音声の送受信が可能になってきたといえる．

そこで，我々が開発した協調 VR 空間のための音声通話ライブラリ vocAL(VOICE Communication Audio Library) では，音声仕様を 32KHz sampling, 16bit, ステレオとする．プラットフォームとして，IRIX と Linux に対応させた．音声の入出力にはクロスプラットフォームのオーディオ I/O ライブラリである PortAudio[9] を利用した．

また，VR 空間上に協調者として表示されるアバ

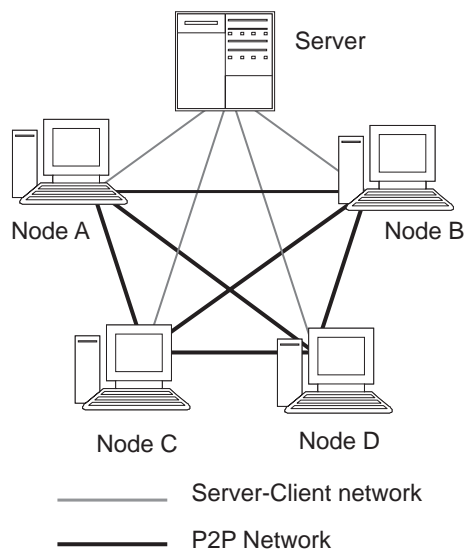


図 2: HybridP2P

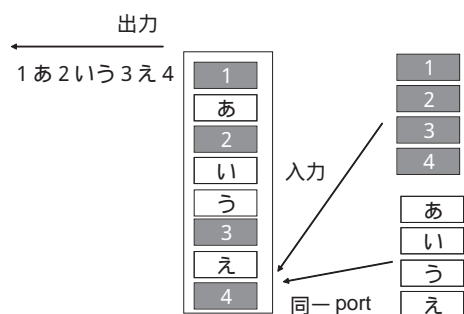


図 3: 同一ポートで音声を受信した場合

タの位置座標は、位置の変化があるたびに座標データが送信される。この座標データを用いて、発話者アバタとの位置関係を計算して PAN と音量を動的に調整する。これによって、音による位置関係や距離感も呈示可能となる。

3.4 Hybrid P2P

3.4.1 音声の多地点同時通信

本研究では、多地点同時通話を可能とするために Hybrid P2P 型のネットワークを構築し、音声転送用プロトコルとして UDP を使用した。この場合、全ての通信路に対して、単一のポートを割り当てると同時に多数のユーザが会話をを行った場合、音声混信の問題が生じる。また UDP で通信を行っているため、図 3 のように、常に順番良く相手からの

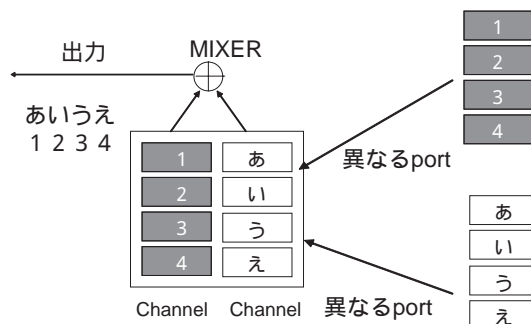


図 4: 異なるポートで音声を受信した場合

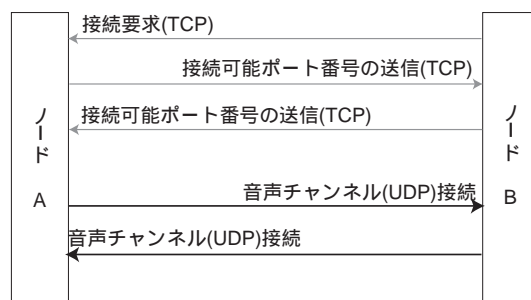


図 5: 音声通信経路の確立

データを受信しているとは限らない。そこで、図 4 に示すように、vocAL では 1 つの通信路に対して 1 つのポートを割り当てることによって、これらの問題を解決した。よって、 n 個のノードと通話するには n 個のポートを確保する必要がある。

図 5 で、一つの通信路についての通信手順を説明する。まず、1 つのノード (ノード A) が立ち上がっているとして、新しいノード (ノード B) が接続する。この時ノード B は接続先であるノード A の IP アドレスを、あらかじめサーバから受信して知っておく必要がある。次に、ノード A は接続してきたノード B に、現在未使用状態となっているポート番号を TCP で通知する。ノード B はそのポート番号を受信したら、自分の空いているポート番号をノード A に通知する。この通知を受けて、ノード A はノード B が使用するポート番号を使用状態に変更する。お互い通知されたポート番号に対して UDP で接続する。なお、接続状況の確認は TCP を使って行う。ノード B が切断した場合は、ノード A はノード B が使用していたポート番号を未使用状態に変更する。

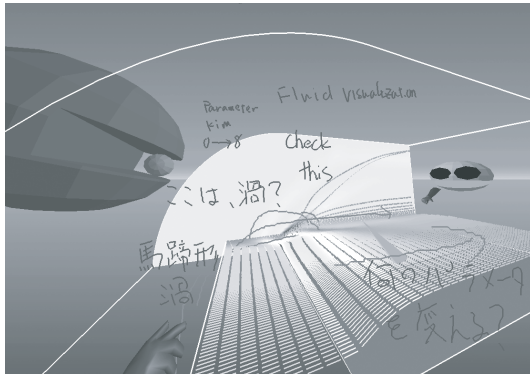


図 6: 協調可視化 VR 環境

4 CVEでのコミュニケーション

4.1 PDA アプリケーションによる注釈描画

我々は、共有 VR 空間に PDA を用いた 2 種類の注釈付与手法を実装している。1 つは PDA のボタンを押しながら PDA を動かして描画する方法であり、PDA のタッチスクリーンに描画した後 VR 空間に転送する方法である。これらの注釈付与の機能を用いて文字情報によるコミュニケーションが実現できる。

これらの手法はフォントを必要とせず、文字を描くので多言語の文字に対応できる。また、空間に描画された文字は、後日のオフラインでの作業に役立つために保存することが可能である。

4.2 音声通話による協調作業

VR 空間での音像提示には 3 次元音響の応用が考えられる。手段の一つとして、Roland 社の RSS-10 が用いられるが、1 音源に対して 1 台の RSS-10 が必要なので、再生する音源数に応じてコストがかかる [10]。また、双方向の音声通話システムで多数のスピーカを用いると、スピーカとマイクを結ぶエコー経路の数が増加し、コミュニケーションに支障が生じる。これを回避するためには多数のエコーキャンセラが必要となり、システムの複雑化とコスト増大の問題が生じる [11]。

この問題に対して、本研究で構築したシステムでは 1 セットの PC 用スピーカとマイクのみを用いる。ネットワーク環境として、学内 LAN(帯域 100Mbps)

に SGI Onyx と Linux PC を用いて 2 地点を接続して実験を行った。協調作業の様子を図 6 に示す。Linux PC 側に IPT を接続して VR 可視化空間を形成し、Onyx 上ではデスクトップ上でアプリケーションを実行して協調環境を共有する。

協調作業者は VR 空間にアバタとして表示され、手や頭を用いたノンバーバルコミュニケーションと注釈や音声通話を用いたバーバルコミュニケーションを同時に行う。音声やアバタの動きに関するデータは UDP で通信を行い、可視化のパラメータ情報や可視化空間への注釈付与は TCP で通信する。接続状況は、socket ステータスの確認によって行われる。

音声通話においては若干の遅延は生じるが、コミュニケーション自体に大きな支障は見られなかった。また、発話者となるアバタとの位置関係による音量と PAN の動的な調整によって、アバタが左から右へ移動しながら発話すると、音声もアバタの動きに合わせて音像が左から右へ移動する。また、発話しながら遠方へ移動すれば音量も徐々に小さくなることを確認した。

次に、研究室内の Linux マシン 3 台を接続して、多地点双方向での実験を行った。先の実験と同程度の遅延が発生するが、双方向での通話を行うことが確認できた。ローカルマシンでフィードバックによる再生を行った実験でも最低約 0.3 秒の遅延が生じる。これは、音声を取り込む時点で発生していると考えられる。今後は遅延の問題解決と、下記に示す項目に着目した評価実験を検討している。

1. 音質に関する評価

電話レベルの音質との比較実験を行い、音質によるコミュニケーションの違いを定量化する。

2. 文字と音声によるコミュニケーションの評価

文字描画のみ、音声のみ、および両方を組み合わせたコミュニケーションの比較を行い、それぞれの有効な利用法について評価を行う。

3. 発話者の位置関係と音量の関係の評価

今回の vocAL では試行錯誤によって音量と発話者の距離の関係を求めた。今後はより厳密に発話者と聞く位置の関係を求め音による距離感を呈示可能にする

4. 音声圧縮

伝送情報量を小さくするため音声圧縮の技術の応用を検討する。

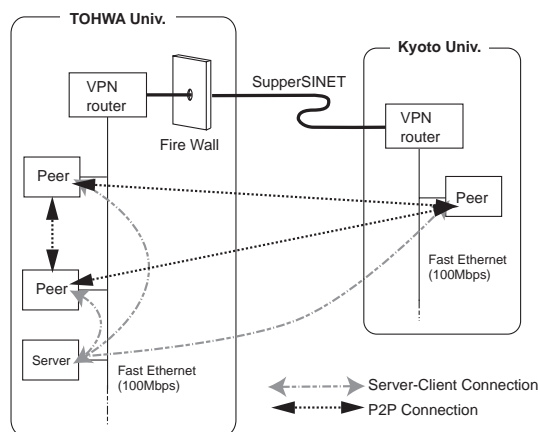


図 7: 協調可視化 VR 環境

5 遠隔地間における CVE

我々は、広域ネットワーク網を通じて円滑な協調作業が可能となる協調 VR 環境の構築を検討している。そこで、今回提案した音声通話ライブラリ vocAL を用いて遠隔地の協調者とのコミュニケーションについて検証するため、実際に遠隔地間での実験を行う。実験は、東和大学(福岡市)と京都大学(京都市)が SuperSINET を介して接続された環境で行う。図 7 にネットワーク構成を示す。接続において、東和大学側に Firewall が設定されているため、京都大学から直接アクセスできない。そこで、VTun(Virtual Tunnel) [12] という高価な VPN 接続装置を用いずに VPN 構築が可能な無償のソフトウェアを利用して、双方の大学からアクセス可能な構成にしている。

今後は、このネットワーク構成でも同様の実験を行う予定であり、また接続拠点を増やすことにより、多地点同時通話の実験評価も行う。

6 まとめ

本稿では、協調 VR 環境におけるコミュニケーション支援に向けた明瞭な音声で多地点間の双方向通話を実現する音声通話ライブラリ vocAL を提案し、その概要について述べた。vocAL は VR 空間で発話者の位置情報によって PAN と音量を調整し、左右の方向感と距離感の呈示が行える。

また、多地点を接続するために HybridP2P によるネットワークを構築し、ローカル環境での実験を

行った結果、遅延が発生する問題があるが、多地点間で明瞭な音声通話を実現した。

今後は遅延の問題について改善を行い、また遠隔地間での高音質な多地点同時通話の実現を目指す。

参考文献

- [1] 廣瀬通孝, 小木哲朗, 玉川憲, 山田俊郎, "没入型コミュニケーションのための高臨場感ビデオアバ", ヒューマンインタフェース学会論文誌, Vol.2, No.2, pp.55-62, 2000.
- [2] Leigh, J., Johnson, A., Brown, M., Sandin, D., DeFanti, T., "Visualization in Teleimmersive Environments", In IEEE Computer, pp. 66-73, 1999.
- [3] Dam, A.V., Forsberg, A.S., Laidlaw, D.H., LaViola, J.J., Simpson, R.M., "Immersive VR for Scientific Visualization: A Progress Report", IEEE Computer Graphics, pp.26-52, 2000.
- [4] Nishimura, K., Abe K., Ishikawa S., Tsutsumi S., Aburatani H., Hirota K., and Hirose M., "Virtual Environment Design Guidelines for Gene Expression Analysis: The Utility of a Lab Bench Metaphor and a Road Metaphor", IEEE Virtual Reality Conference 2004, pp.247-248, 2004.
- [5] <http://www.accessgrid.org>
- [6] <http://www.itua.jp>
- [7] Kukimoto, N., Furusho, Y., Nonaka, J., Koyamada, K., Kanazawa, M., "PDA-based Visualization Control and Annotation Interface for Virtual Environment", In Proc. VIIP 2003, The 3rd IASTED International Conference on Visualization, Imaging, and Image Processing, pp.975-980, 2003.
- [8] Yang, B., Garcia-Molina, H., "Comparing Hybrid Peer-to-Peer Systems", The VLDB Journal, pp.561-570, 2001.
- [9] <http://www.portaudio.com/>
- [10] 久木元伸如, Ewe Chin Huar, 竹田仰, "プロジェクト型没入ディスプレイにおける 3 次元音場生成のための実用的検討", 電子情報通信学会技術報告書 EA99-32, p29-36, 1999.
- [11] 馬屋原立祐, 穂刈治英, 島田正治, "スペクトル補間を用いたステレオステージ制御方法-多地点通信のための新しい要素技術-", 信学会, A Vol. J85-A, No.8, pp.819-832, 2002.
- [12] <http://vtun.sourceforge.net/>