

4 者間討議空間システムの構築

鈴木健治^{*1}, 田中健二^{*1}, 鈴木龍太郎^{*2}, 荒川佳樹^{*1}, 小木哲朗^{*3}

Construction of "4-Point Discussion Space System"

Kenji SUZUKI^{*1}, Kenji TANAKA^{*1}, Ryutaro SUZUKI^{*2}, Yoshiki ARAKAWA^{*1} and Tetsuro OGI^{*3}

Abstract - 4-Point Discussion Space System, a virtual space which connects four remote institutes, has been constructed at Communications Research Laboratory (CRL) as part of the Multimedia Virtual Laboratory (MVL) project. In this system CRL, the University of Tokyo, Tsukuba University and National Institute of Multimedia Education (NIME) are connected with 75Mbps ATM networks. By reproducing eye contact and voice directions, the constructed virtual space presents an environment where four people in four different institutes can have natural conversation with one another as if they are in the same room.

Keywords : Multimedia, Virtual laboratory, Teleconference, ATM

1. はじめに

郵政省通信総合研究所（CRL）では、平成9年度から5カ年計画のマルチメディア・バーチャル・ラボラトリー（MVL）プロジェクトにおいて、3D実画像と3D-CGを用いた非常に臨場感の高いコミュニケーション環境の実現を目指している。3D情報空間“UNified Virtual EnviRonment & Space（UNIVERS）”の構築とこれらの情報空間を高速ネットワーク上で共有する研究を行ない[1]、遠く離れた研究所を結び、共同研究や共同実験を行うネットワーク上の仮想研究所を構築しようとしている。その中で実画像をベースに4者があたかも同じ場所に会して議論しているかのような、“4者間討議空間システム”を実現した。これは、CRL/UNIVERS 東京大学インテリジェント・モデリング・ラボラトリー（IML）/CABIN[2],[3] 筑波大学ベンチャービジネスラボラトリー（VBL）/Co CABIN メディア教育開発センター（NIME）の4地点を75Mbps ATM回線で結び、各地点間の視線と音声到来方向を再現し自然な会話環境を提供するものである。本報告ではそのシステムの概要と評価について述べる。

高臨場感TV会議システム等の研究[4]は数多く報告されているが、“より自然な対話環境の追求”ということに主眼を置いて、特に次の点を考慮して通信システムを構築した。

お互いに相手があたかも同一空間内において対話しているかのような環境として、重要な条件として、

- 1) 会話における音声遅延時間が十分短いこと、
- 2) 音声到来方向が確保されていること、
- 3) 相手映像からその表情、口の動きが見てとれること、
- 4) 言葉の理解度を高めるためのリップシンクがとれていること、

5) 会話時に相手映像の視線方向が確保され、アイコンタクトが得られること、

等が挙げられる。これらのことから、それぞれ3画面に各地点の実画像を表示するため、各地点に3本のカメラとマイクを用意し、CRLをハブとした75Mbpsの広帯域なATM回線を介して、低遅延NTSCコーデックを用いて実験を行った。

2. 実験方法

2.1 ATMネットワーク

図1にATMネットワーク構成を示す。CRL所内は155Mbpsで繋ぎ、東京大学IML/CABIN、筑波大学VBL Co-CABIN、NIME間を、75Mbps ATM回線で接続。低遅延コーデック間及び、CRL/Onyx2、東京大学Onyx2、筑波大学Onyx等の計算機間のIP/ATMに、PVCでチャンネルを張って実験を行った。

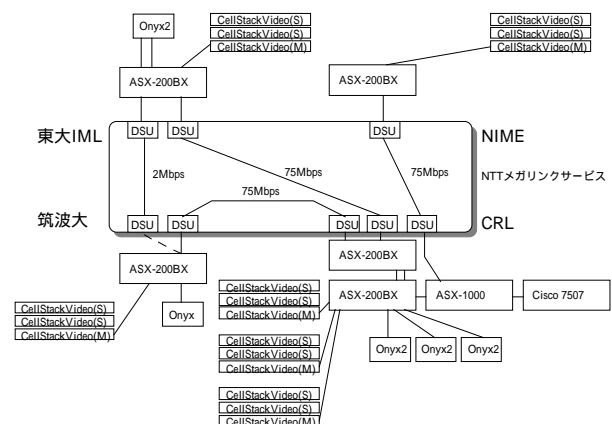


図1 ATMネットワーク構成図
Fig. 1 ATM Network configuration.

2.2 低遅延NTSCコーデック

ATMインターフェイスを持つCellStackVideo(M)マスター機にCellStackVideo(S)スレーブ機2台を専用ケーブル

^{*1}: 郵政省通信総合研究所, ^{*2}: 通信・放送機構,
^{*3}: 東京大学インテリジェント・モデリング・ラボラトリー
^{*1}: Communications Research Laboratory, M.P.T., ^{*2}: NeLS,
TAO, ^{*3}: IML, University of Tokyo

でスタック状に接続し、NTSCビデオ信号3ch、ステレオ音声3chの双方向通信が可能となる。TV会議モードにおいて画像は、720×240 pixels 30f/s NTSC Motion JPEGで、音声は16bitPCM 48kHzサンプリングとなっており、ATM帯域としては(画像10~15Mbps+音声1.73Mbps)×3=35~50Mbpsとなる。CellStackVideoのATM PVCの設定は、マスター機のAudio: VPI=0, VCI=200; Video: VPI=0, VCI=201, スレーブ機-1 Audio: VPI=0, VCI=202; Video: VPI=0, VCI=203, スレーブ機-2 Audio: VPI=0, VCI=204; Video: VPI=0, VCI=205となる。

2.3 CABIN及び、UNIVERS他の主要諸元

CABIN(図2)は、2.5m×2.5mの高精細立体表示スクリーン上下左右前の5面で構成され、映像を見渡す際の視野角が大きく実空間に近い視空間が生成される。CG/3D-CG映像を見回す際の視点移動範囲も広い。また、スイッチを切り替えれば、NTSC信号の入力を直接CABIN内に投影することができる。ここでは左右前の3面のみを使用した。UNIVERSは、1画面100インチのリア投影型高精細プロジェクタ×3台と、それらにCG又は、実画像を処理して出力するためのグラフィックワークステーションOnyx2×3台と周辺機器で構成される。UNIVERSの外観を図3に示す。特長として、スクリーン前の人物をカメラで撮影するのに十分な光量であっても、スクリーンに写る映像が見にくくならないよう、レンチキュラー面にブラッ

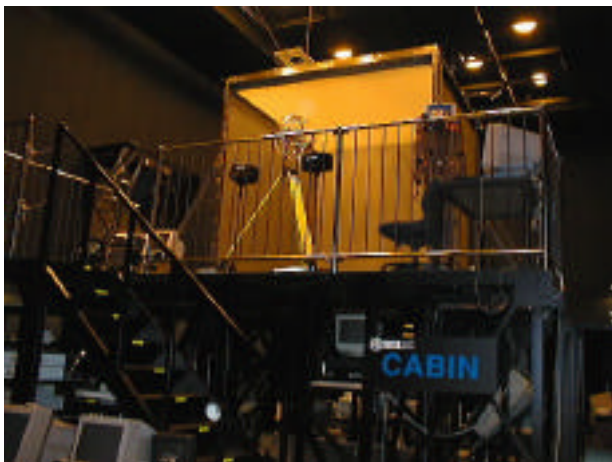


図2 CABINの外観

Fig. 2 Immersive multiscreen display "CABIN".



図3 UNIVERSの外観

Fig. 3 Presence 3-display "UNIVERS".

クストライプを成型したスクリーンを採用している。中央のスクリーン装置は固定で、左右のスクリーン装置は用途に応じて角度を調整できるようになっている。液晶シャッター方式による立体視ができる。筑波大学VBL/Co CABINは、1画面90インチのプロジェクタ×3台で左右前の3画面で構成されている。NIMEにおいては、3m×3m5.5面の没入型仮想空間システム[5]TEELeXを開発中で、今回は40インチのプラズマディスプレイ3台を並べて、左右前の3画面で構成された。

2.4 カメラ、音響機器

4者間討議空間実験用のカメラはマイクに取り付け、なるべくカメラを意識させないために小型のスーパーマイクロカメラを選んだ。

主要諸元は、

41万画素1/4インチ高精細CCD

有効画素：水平768画素 垂直494画素

有効撮像面積：水平3.65mm 垂直2.74mm

解像度：水平：470TV本以上 垂直：350TV本以上

レンズ：4mm 水平方向視野角51度

である。マイクは鋭指向性コンデンサマイクロフォンを使用し、再生系は、L、Rのステレオ再生としている。エコーキャンセラーは無く、ボリュームマトリックスミキサーの音量調整によって最適化している。

2.5 4者間討議空間システムの構成

4者間討議空間システム概念図を図4示す。水平方向360度の空間を4者で90度ずつに分割してお互いに共有する。相手方3者に対してそのスクリーンの前に置いたそれ

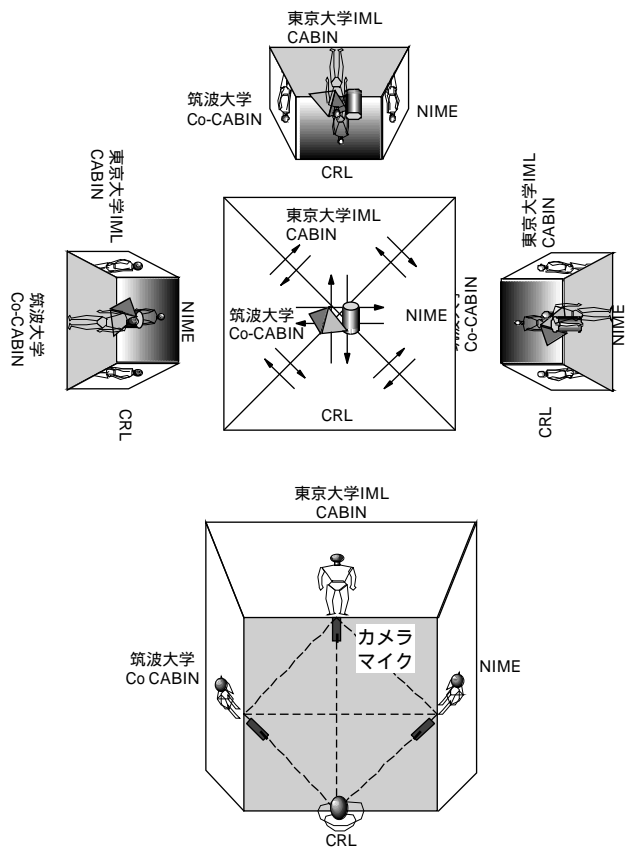


図4 4者間討議空間システム概念図

Fig. 4 4-Point Discussion Space System.

ぞれ別のマイク（3本）で集音しその相手のみに送る事によって、マイクに向かって小声で話せば、他の2者には伝わりにくく私語も可能としている．再生側では、左右2つのスピーカに音圧レベル調整によって左(L),中央(C),右(R)に定位させる．

4者を各地点それぞれ正方形の辺の中央に位置するように調整し、3本のカメラを3面スクリーンのそれぞれ中央部に配置し、画面の中央部に人物の顔が入るように撮影してその相手方に送信する．この時中央(C)のカメラは中央のスクリーンから人物を正面から捉え、他の左右のカメラは人物を45度の角度に捉えることになる．この配置によって人物の顔が左,中央,右のスクリーン中央を望んだとき、45度ずつの角度差が生じる．その映像はそれぞれ相手先のスクリーンに映し出された時、お互いの位置関係が保存され人物がいずれの方向を向いているかが判断できる．

図5にシステム構成図を示す．CRL 筑波大学VBL, CRL 東京大学IML, CRL NIME間のATM75Mbpsの回線を通してそれぞれ対向する低遅延コーデック×3台で接続し、音声と映像をやり取りする．4地点の3chの音声信号は、一旦ボリュームマトリクス音声ミキサーに入力し、

- CRL(L) 筑波(R)
- CRL(C) 東大(L) + 東大(R)
- CRL(R) NIME(L)
- 筑波(L) 東大(R)
- 筑波(C) NIME(L) + NIME(R)
- 筑波(R) CRL(L)
- 東大(L) NIME(R)
- 東大(C) CRL(L) + CRL(R)
- 東大(R) 筑波(L)
- NIME(L) CRL(R)
- NIME(C) 筑波(L) + 筑波(R)
- NIME(R) 東大(L)

と音声レベル調整して振り分け、それぞれの地点で左

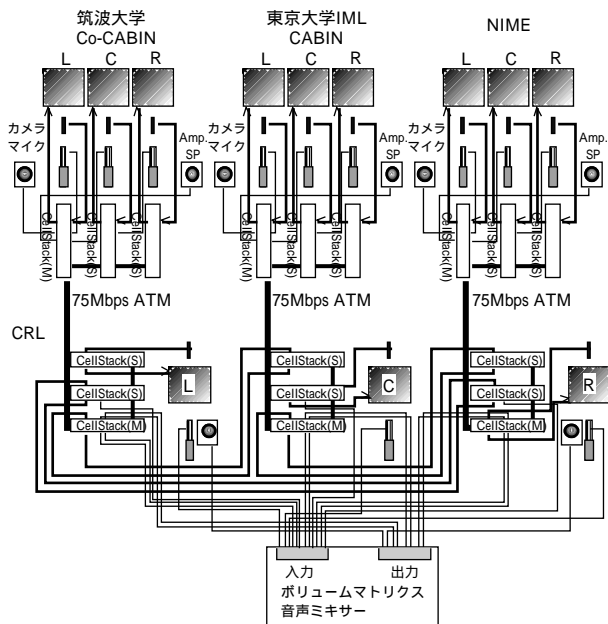


図5 4者間討議空間システム構成図
Fig. 5 System configuration.

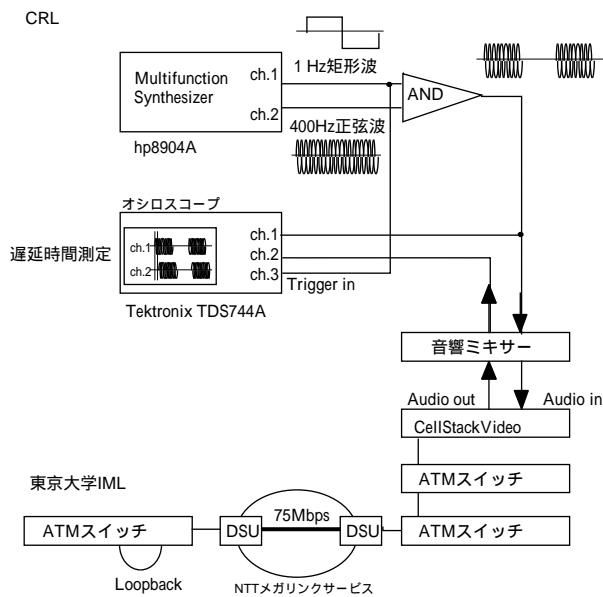


図6 音声遅延時間測定
Fig. 6 Measurement of audio delay time.

(L), 右(R), 中央(C) に定位させる．

4地点の3画面の映像信号は、それぞれデコード後エンコードして必要な所へ再送信するようになっているが、実際はATMスイッチ内で映像信号のPVCを交換した．

- 筑波VPI=3, VCI=201 東大VPI=4, VCI=205
- 筑波VPI=3, VCI=203 NIME VPI=2, VCI=203
- 東大VPI=4, VCI=201 NIME VPI=2, VCI=205

2. 6 低遅延NTSCコーデックの音声遅延時間測定

図6に低遅延NTSCコーデックの音声遅延時間測定装置構成を示す．1秒毎に発せられる400Hzの音声信号を図のように低遅延NTSCコーデックに入力しエンコードして相手先ATMスイッチで折り返す．再び低遅延NTSCコーデックでデコードし生音声信号との遅延時間をシンクロスコープで測定する．片道の伝送経路に比べて伝送部分が2重になってしまうが、CRL 東京大学間回線距離27kmから光ファイバ内伝送時間は0.135msecであり、コーデックでの遅延量に比べて小さく支配的でない．

3 実験結果と考察

高速なATMネットワークを使って高臨場感表示システム間を、低遅延なNTSC実画像・ステレオ音声で接続し、水平方向360度を90度ずつ分割した4者間討議空間システムを実現した．お互いの視線方向、音声到来方向も確保し、今までの多地点TV会議システムには無い自然な対話環境を再現できた．図7、図8に東京大学IML/CABINで撮影した実験風景を示す．図7では左に見えるスクリーンが中央でCRL、右スクリーンに筑波大学が写っている．同じく図8では、左にNIME、中央にCRLが写っている．

CRLと東京大学IML間の低遅延NTSCコーデックによる片道の伝送遅延時間は、約16msecと十分小さく意志の疎通を図るのに十分であった．16msec程度であれば、空気中の音速が約340m/sであるため例えば、2者が約5.4m離れて会話している場合に生じている遅延時間に相当し、殆ど無視



図7 4者間討議空間実験-1
Fig. 7 Four-Locational Discussion Space experiment-1



図8 4者間討議空間実験-2
Fig. 8 Four-Locational Discussion Space experiment-2

できる範囲内である。

CABIN においてはスピーカを入り口部分の左右に置いたが、左右の音の分離と中央の人物からの音声は映像に引き込まれる形でその音声到来方向が確保されるのが確認された。

低遅延NTSCコーデックのビデオ会議モードにおける画質が720×240pixels, 30fpsと、水平解像度が240本であり100インチを越える大画面のプロジェクトに投影した場合十分な画質が得られないという問題もあったが、相手映像からおよその表情の変化は見て取れた。また、音声と映像の同期は取れていて、ずれを知覚することはなかったが、音量調整が不十分で、音量が小さいために聞き取りにくい場合に、相手の口の動きから言葉を補間できるほどの解像度はなかった。スクリーン中央に配置したカメラ・マイクが目障りではあったが、会話時に相手とのアイコンタクトが得られ、別の3者がそれぞれ対話する場合においてのそれぞれの視線方向の一致がみられ誰と誰が会話しているかがすぐ分かった。これは視線方向に限らず指などで指し示した場合にも成立した。

4 おわりに

今後は、この4者間討議空間内に議論対象となるモデルや討議資料を実画像に実画像又はCG, 3D-CG画像で融合し、さらに共有ボード等を4者がお互いに違和感なく利用できる設定を検討する。次に、人物を切り出し背景を実画像又はCG, 3D-CG画像で共有し、より一体感を表現する。さらに、現在スクリーンの真ん中に配置している目障りなカメラを例えば、正面スクリーンの4隅に配置し、人物を実画像からCG画像の正面顔に変換するバーチャルカメラにして、より自然な共有空間を追求する予定である。

本研究は、マルチメディア・バーチャルラボラトリーを構成する上で重要となるネットワーク部分の基礎的技術の開発となり、CABIN, Co CABIN, UNIVERS等の異空間を接続しコミュニケーションの基本となる要素を明らかにするものとして期待される。

謝辞

日頃から協力頂いているMVL 開発推進協議会、郵政省通信政策局及び、研究協力頂いた東京大学IML、筑波大学葛岡英明氏、NIME大西仁氏を始め関係各位に感謝致します。また、ネットワーク構築を行う上で色々ご助言頂いた、NTT松岡英晃氏、NKエクサセブ野修一氏に感謝致します。なお、本研究は科学技術庁の国内留学制度によって、東京大学IMLに留学中に行ったものである。

参考文献

- [1]鈴木健治, 田中健二, 鈴木龍太郎, 荒川佳樹, “UNIVERSを用いたネットワーク実験計画”, 信学会ソサイエティ大会, A-16-8, 1998-09.
- [2]廣瀬通孝, 小木哲朗, 石綿昌平, 山田俊郎, “没入型多面ディスプレイ(CABIN)の開発”, パーチャルリアリティ学会大会論文集, Vol.2, pp.137-140, Sep 1997.
- [3]廣瀬通孝, 小木哲朗, 石綿昌平, 山田俊郎, “多面全天周ディスプレイ(CABIN)の開発とその特性評価”, 信学論(D-II), Vol.J81-D-II, No.5, pp.888-896, May 1998.
- [4]並木郁夫, 目黒義隆, 青木茂明, 入江一成, 野村知義, 斉藤一重, “高臨場感マルチメディア通信会議システムの構築と評価”, 信学論(B-I), Vol.J80-B-I, No.6, pp.338-347, Jun 1997.
- [5]杉本裕二, 浅井紀久夫, 大澤範高, “高輝度没入型仮想空間システムの開発”, 信学会総合大会, A-16-17, 1999-03.