

未来の電話 t-Room: 対面環境の再現を目指すビデオコミュニケーションシステム

Future Phone t-Room:

Video Communication System Towards Reproduction of Face-to-Face Environment

平田 圭二
Keiji Hirata

梶 克彦
Katsuhiko Kaji

NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

1. まえがき

現在我々はビデオ共有面を持つビデオコミュニケーションシステム t-Room を開発している [3,2,18]. t-Room の目標は、時間的に距離的に離れた空間にいるユーザとあたかも実対面しているような協調コミュニケーション環境を実現し、コンテンツ創出のためのメディアを構築することである。

従来のビデオ会議システムの多くはユーザが視聴する映像と音をできるだけ忠実に対面環境のそれと等しくすることを目指している。その要件は3つあるだろう：1) 高画質・高音質・低遅延を達成する、2) ユーザの周囲の作業情景や物理的位置関係（空間キュー）をできるだけ大面積、広角度で再現する、3) 遠隔地のユーザとの間で映像・音情報が対称的であること。特に 3) の映像情報の対称性に関する性質は WYSIWIS (What you see is what I see) [13] あるいは reciprocity of perspective [1] と呼ばれ、この対称性が成立しない場合を presence-disparity problem [15] という。2), 3) に主眼を置いたシステムとして Hydra [12], MAJIC [11], Gaze-2 [17], MultiView [9,10] 等がある。

一方、我々は次の4点を従来のビデオ会議システムが抱える問題点と考えている：a) 遠隔地の空間とローカルの空間の間を自由に移動できない、b) スクリーンに表示された物体の像を指差しても遠隔地のユーザに伝わらない、c) 3 地点以上を接続すると対面環境のような空間キューを保存するのが難しい、d) 記録された映像を picture-in-picture で再生すると再生映像が入れ子になってしまい、任意回の記録と再生の繰り返しができない。これら a)~d) の問題点を解決する手段としてビデオ共有面があり、それを導入した遠隔協調作業システムとして VideoDraw [16], ClearBoard [4], Agora [8], VideoArms [15] 等がある。

前述の 2), 3) の要件を満たし、a)~d) の問題を解決するために、我々は、大型液晶ディスプレイ (LCD) でビデオ共有面を実現し、その LCD でユーザ空間を囲い込み、そこに実物大で遠隔地の様子を投影する方式を提案した [3,2]。この閉じたユーザ空間を囲む LCD に、2 か所以上の遠隔地の映像を重ね合わせて投影することで、遠隔地とローカルのユーザ空間全体が重ね合わされ、移動する人や物の間に成立している空間キューを遠隔地との間で保存することができる。さらにユーザ空間をまるごと記録しそのまま（重ね合わせて）再生すれば、記録と再生を任意回繰り返すことができる。

本方式では、遠隔地の映像・音と記録された過去の映像・音を同格に扱うことができるようになった。これより、現在から再生中の過去の活動を参照できるので（電子メールやブログにおける引用に相当する）、時間の壁を超えた非同期インタラクションが可能になる [14]。非同期インタラクションを実現している既存メディアには、電子メール、ブログ、留守番電話、動画投稿サイト等があるが、いずれもテキスト主体であり映像や音に関しては限定されていた。

2. 背面スクリーン囲い込み方式

図 1 に、我々が提案した実対面環境を遠隔地との間で再現するための方式を示す。実インタラクション空間を人 A, B, C のそれぞれに対して複製し、ユーザを囲い込む背面のスクリーンに遠隔地のユーザ像を投影する。図中の各空間（部屋）では、ユーザ空間を取り囲むようにスクリーンとカメラが交互に3台ずつ配置され、ユーザはスクリーンの直ぐ前面に位置する。☐ の機能は、対向するスクリーン前面の物体の映像だけを抽出し、スクリーンに映っている映像をキャンセルすることである（映像エコーキャンセラ）。⊕ の機能は、2 つ以上の遠隔地からの映像を正しい位置に正確に重ね合わせることである（オーバーレイヤ）。このようにして例えば、部屋 2 で B を撮影している

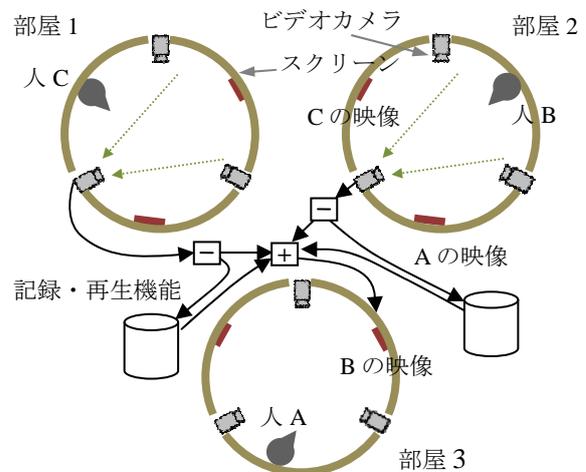


図 1: 空間複製とユーザを囲い込む背面スクリーンへの映像投影

カメラの映像が部屋 1, 3 に転送される。映像エコーキャンセラ出力は転送されると同時に記録され、後で再生する時に読み出されオーバーレイに転送される。図 1 では簡単のため必要な配線のみ描いた。

3. t-Room におけるアプリケーション例

ビデオコミュニケーションシステム t-Room は、図 1 の方式に基づいて実装されている。離れた時間と空間の区別なく対面環境を自由に再現できることを利用した 4 つのアプリケーションを紹介する[5]。ここで用いられている t-Room 1 セットは、65 インチ液晶ディスプレイ 8 台と HD 画質ビデオカメラ 8 台を 10 角形に配置した構成をとっている。

リモコン制御の映像再生

対面環境を再現する例として、過去に録画した映像をリモコン操作で再生している様子を図 2 に示す[6]。上図が最初に録画された過去の映像である。下図は、スクリーンに表示されている 7 つのオブジェクトを示している：1) 現在の遠隔地ユーザ、2) 録画されたユーザ、3) 録画シーンの再生を制御するための操作ボタン群（リモコンでクリックする）、4) ビデオプレイヤ表示ウィンドウ、5) ビデオプレイヤ操作ボタン群、6) 現在のローカルユーザが操作しているリモコンのカーソル、7) 現在の遠隔地ユーザが操作しているリモコンのカーソル。ローカルや遠隔地の区別なく、対面で同室にいるかのようにビデオプレイヤを操作できる。

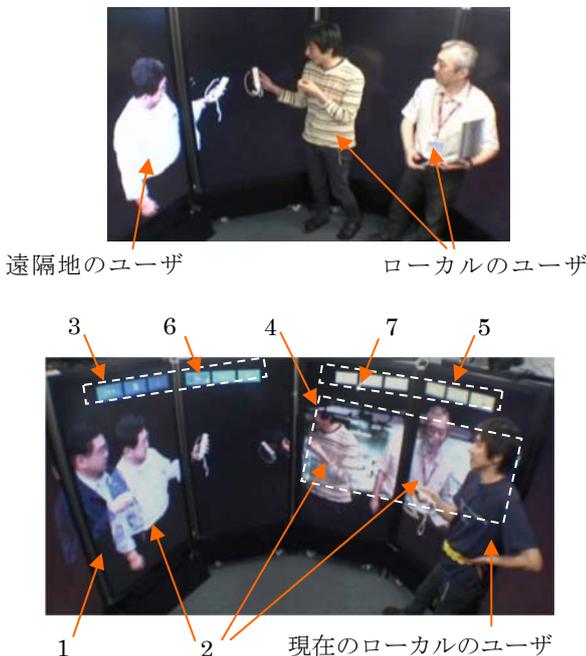


図 2: 録画されたシーン（上）。リモコン制御で再生される様子（下）

シーン切替え時のエレベータ効果

t-Room どうしの接続や背景映像などを変更する際、透明なエレベータに乗って上昇・下降するようにシーンを切り替える効果を実装した（図 3）[7]。対面環境を再現する際



図 3: エレベータが下降するようなシーン移動

の効果として有効であることが確認された。

遠隔ゴルフレッスン

図 4 は、時間と空間を超えた対面環境の再現を利用して、ローカルの生徒が遠隔地のレッスンプロからゴルフスイングのレッスンを受けている様子である[7]。レッスンプロと生徒のジェスチャ、指差し、体躯の移動等 t-Room 内で起きる全ての事象が記録され、その時に再生されていた過去の映像も記録され、後で同様に再生することができる。遠隔地のレッスンプロから見ると、現在の生徒も過去の生徒も区別することなく、あたかも対面しているようにそのスイングフォームを点検できる。

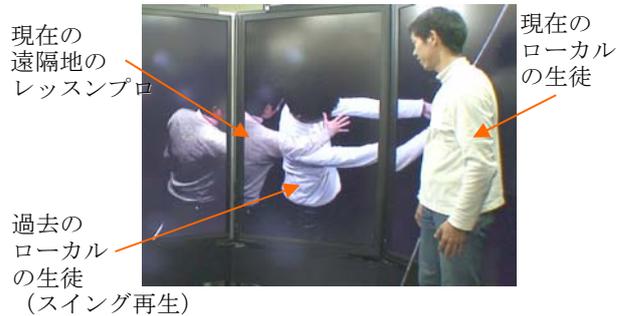


図 4: 生徒のスイングを再生しながら遠隔地のレッスンプロがフォームの修正点を指摘

非同期音楽セッション

時間と空間を超えて対面環境を再現する一例である。従来のマルチトラック多重録音では音のみを扱うことができた。t-Room の記録・再生機能を用いると、マルチトラック多重録音と同様の手順で、それまで録画したパートを再生しながら、あたかも対面環境でバンド演奏しているかのように、新たなパートを順番に録画していくことができる（図 5）。上図は、すでに録画済みのドラムス (Dr)、ベース (B)、キーボード (Kb) を再生しながらギター演奏の追加録画を行っているところである。下図は、上図で録画された各楽器パートの内、B の演奏だけを入れ替えるために、新たに B だけ演奏し直しているところである。

4. おわりに

本稿では、時間的に距離的に離れた空間にいるユーザとあたかも実対面しているような協調コミュニケーション環

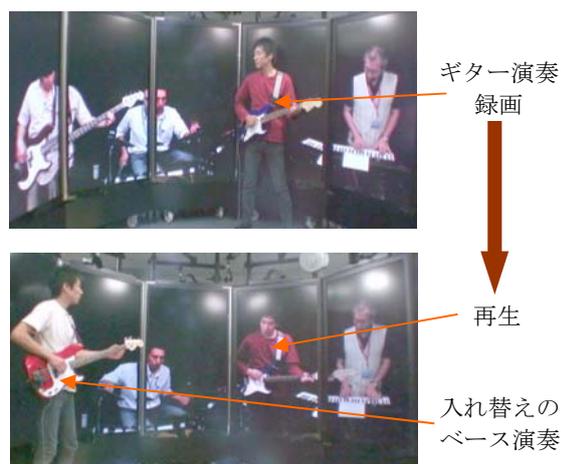


図 5: Dr, B, Kb 再生をバックにギター演奏を録画 (上). 前の B 演奏を入れ替えるために新たに B を演奏 (下)

境を実現する方式と、その実装である t-Room 上のアプリケーションを 4 件紹介した。これらアプリケーションにおいて、ユーザは従来のビデオ会議システムより実対面に近い環境で遠隔協調作業を遂行することが期待される。現在、この点を検証するため、被験者を用いた遠隔協調作業実験を実施中である。この実験を通じて、presence-disparity problem への対処法、ユーザの座席配置や空間キューの伝達が協調作業に及ぼす影響等を調査している。

我々の周囲にはスポーツ、料理、工芸など非言語的なスキルを必要とするような領域が多数あり、そのようなスキルを利用したコンテンツ創出の場面が多数ある。我々が提案した対面環境を再現する方式は、従来のテキスト主体のコミュニケーションメディアを汎用かつ効率的に補完できる可能性があり、そのようなコンテンツ創出の環境として適しているのではないかと考えている。今後は電話のように日常生活のあらゆる場面で用いられるようなコミュニケーションメディアへと t-Room を発展させていきたい。

参考文献

1. Heath, C. and Hindmarsh, J. Configuring Action in Objects: From Mutual Space to Media Space. *Mind, Culture, and Activity*, 7(1&2), 81-104 (2000).
2. Hirata, K., Harada, Y., Takada, T., Aoyagi, S., Shirai, Y., Yamashita, N., Kaji, K., Yamato, J., and Nakazawa, K. t-Room: Next Generation Video Communication System. In *Proceedings of World Telecommunications Congress 2008* (with IEEE Globecom 2008).
3. Hirata, K., Harada, Y., Takada, T., Yamashita, N., Aoyagi, S., Shirai, Y., Kaji, K., Yamato, J., and Nakazawa, K. Video Communication System Supporting Spatial Cues of Mobile Users. In *Proceedings of CollabTech 2008*, IPSJ, pp.122-127.
4. Ishii, H., Kobayashi, M., and Arita, K. Iterative Design of Seamless Collaboration Media. *Communications of the ACM*, 37, 8 (August, 1994), 84-97.

5. 梶克彦, 平田圭二. t-Roomにおけるデスクトップ画面共有機能の実装とその応用. 日本ソフトウェア科学会第 25 回大会論文集, 2B-2 (2008).
6. 梶克彦, 平田圭二, 長尾確. t-Roomにおけるマルチユーザ・マルチロケーションポインティングシステム. 日本ソフトウェア科学会第 25 回大会論文集, 2B-3 (2008).
7. 梶克彦, 平田圭二, 原田康徳. t-Roomの仮想共有面と過去再生機能を用いたアプリケーションについて. 7th Forum on Information Technology (FIT2008), 第 4 分冊, RM-005, pp.51-54, IPSJ/IEICE-ISS (2008).
8. Kuzuoka, H., Yamashita, J., Yamazaki, K., and Yamazaki, A. Agora: A Remote Collaboration System that Enables Mutual Monitoring. In *Proceedings of CHI'99 Extended Abstracts*, 190-191.
9. Nguyen, D. and Canny, J. Multiview: Spatially Faithful Group Video Conferencing. In *Proceedings of CHI2005*, pp. 799-808.
10. Nguyen, D. and Canny, J. Multiview: Improving Trust in Group Video Conferencing Through Spatial Faithfulness. In *Proceedings of CHI2007*, pp. 1465-1474.
11. Okada, K., Maeda, F., Ichikawa, Y., and Matsushita, Y. Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design. In *Proceedings of CSCW'94*, pp. 385-393.
12. Sellen, A. J. Speech Patterns in Video-Mediated Conversations. In *Proceedings of CHI '92*, 49-59.
13. Stefik M., Bobrow D., Foster G., Lanning S., and Tatar D. WYSIWIS Revised: Early Experiences with Multiuser Interfaces. In *Proceedings of CSCW 1987*, pp. 147-167.
14. Takada, T. and Harada, Y. Citation-Capable Video Messages: Overcoming the Time Differences without Losing Interactivity. In *Proceedings of Information Spaces and Visual Interfaces* (2000), 31-38.
15. Tang, A. and Greenberg, S. Supporting Awareness in Mixed Presence Groupware. ACM CHI Workshop on Awareness systems: Known Results, Theory, Concepts and Future Challenges (2005).
16. Tang, J. C. and Minneman, S. L. VideoDraw: A Video Interface for Collaborative Drawing. In *Proceedings of CHI '90*, 313-320.
17. Vertegaal, R., Weevers, I., Sohn, C., and Cheung, C. GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction. In *Proceedings of CHI 2003*.
18. Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H., and Harada, Y. Impact of Seating Positions on Group Video Communication. In *Proceedings of CSCW 2008*, pp.177-186.