

# ビデオアバタを用いた多地点間 CAVE コミュニケーション

酒井 満隆<sup>\*1</sup> 小木 哲朗<sup>\*1</sup> 立山 義祐<sup>\*2</sup> 江原 康生<sup>\*3</sup> 宮地 英生<sup>\*4</sup>

## Communication among multiple CAVE sites using Video Avatar Technology

Mitsutaka Sakai<sup>\*1</sup>, Tetsuro Ogi<sup>\*2</sup>, Yoshisuke Tateyama<sup>\*3</sup>, Yasuo Ebara<sup>\*4</sup> and Hideo Miyachi<sup>\*5</sup>

**Abstract** - Video avatar is technology that represents a person's image in the virtual world. In late years, immersive projection technology, such as CAVE, CABIN had become very popular as a three-dimensional visual display. In this study, video avatar technology was applied to the communication in three-dimensional immersive virtual world. We can perform, in particular, the mutual communication that shared the three-dimensional virtual world through a high performance network.

**Keywords:** Video Avatar, IPT, real-time, network

### 1. はじめに

近年、遠隔地の人々とコミュニケーションをとる多くのが登場してきている。その中で、相手の様子を直接見て取れる手段としてはテレビ電話やビデオ会議システムなどが挙げられる。リアルタイムで映像を送りあうことで、声や文字だけでは伝わりきれない多くの情報を共有することが可能である。しかし、実際その映像のサイズはそれぞれ機器のスクリーンサイズによって制限されている。等身大での通信という点からすれば困難なものといえる。実写映像ではなく自分がアバタとなり分身に置き換わることでコミュニケーションをする方法もある。その1つの手法としてCGを用いたアバタ表現が挙げられる[1]。これにより、その時々相手の身振りや手振りの情報を伝えることができる。

そして、最近では映像の通信としてビデオアバタを用いるという手法も考えられてきた[2]。ビデオアバタとは、ビデオカメラで撮影している映像からリアルタイムで仮想の人物表現を行う技術である。あくまで自分自身の分身であり、実写像であるところにビデオアバタの特徴があると考えられる。

また、3次元立体映像の表現方法としては没入型ディスプレイが開発されてきた。代表的なものとしてはCAVE[3]が挙げられる。ユーザが立方体状のスクリーンに囲まれた環境におかれることで、広視野の没入空間を体験することができる。その没入型ディスプレイ同士をネットワークで

結ぶことで臨場感の高いコミュニケーションを実現することができる[4]。

本研究ではビデオアバタを用いてコミュニケーションを行うと同時に、多地点間の没入型ディスプレイをネットワークで接続して相互通信を行うことを目的としている。実時間でビデオアバタを生成して遠隔地にあるCAVEへ相互に送信し投影する。ここでは特に3地点間のCAVE間ネットワークを確立し、各サイトで同時にCAVEを動かしながらビデオアバタ相互通信を行った。

### 2. ビデオアバタ

#### 2.1 背景差分法

通常ビデオカメラで人物の映像を撮影している場合には、目的としている人物の他に、その空間に配置してある物体や背景領域も含まれる。しかし、ビデオアバタに必要なものはあくまで人物像のみである。それには目的としない余分な領域を除かなくてはならない。人物の抽出方法に関して、代表的なものにブルーバックを用いる方法がある。これは撮影時の背景を青色にし、その青色領域を切り抜く方法である。特定の色のみを切り抜くことから比較的きれいな処理を行うことが可能である。しかし、これには特殊な撮影環境を用意しなければならないといった問題点も考えられる。また、CAVE間での相互通信を行おうとした場合、撮影対象者はユーザその人である。スクリーンには映像が投影されているため、内部にいるユーザを撮影するのにブルーバックのような環境を設置することは困難である。

本研究では、CAVEのユーザ同士が相互通信をする中でビデオアバタとなることが目的である。そこで、背景差分法をもちいることで、全身の実写人物像をリアルタイムで獲得している。背景差分法とはあらかじめ撮影しておいた背景画像と、その背景画像と同じ環境下で撮影した観測映像の差分から目的とする人物を抽出する方法である。

\*1: 筑波大学大学院 システム情報工学研究科

\*2: 東京大学 インテリジェント・モデリング・ラボラトリー

\*3: 京都大学 学術情報メディアセンター

\*4: 株式会社ケイ・ジー・ティー ビジュアリゼーション事業部

\*5: Graduate School of Systems and Information Engineering, University of Tsukuba

\*2: Intelligent Modeling Laboratory, the University of Tokyo.

\*3: Academic Center for Computing and Media Studies, Kyoto University

\*4: Visualization, KGT



図 1 実行例  
Fig.1 example of practice.

図 1 に観測映像のある 1 フレームにおける背景差分処理の例を示す。背景画像と観測映像の RGB 値の絶対値を求め、その値がある閾値より大きい場合のみ描画をする。これにより目的とする人物を切り抜くことができるのである。

### 2.2 3次元仮想世界との合成

生成したビデオアバタは 3 次元仮想世界との合成をして描画をする。3 次元仮想世界の構築には OpenGL を使用してモデリングを行った。

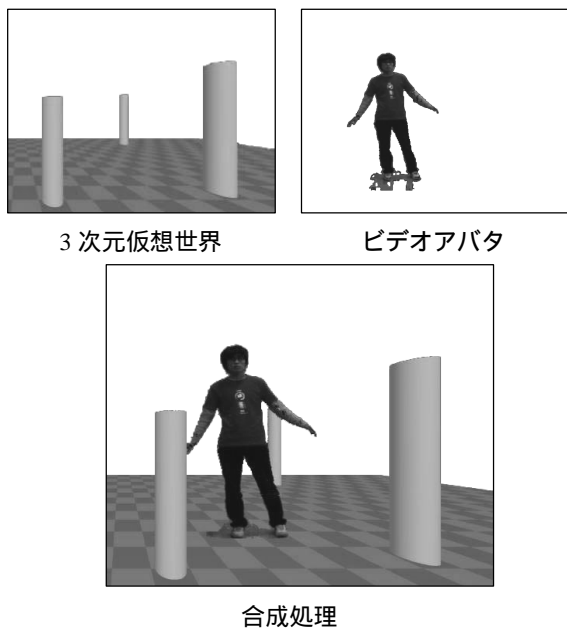


図 2 合成例  
Fig.2 example of composition.

図 3 はその合成例を示したものである。OpenGL で構築した仮想世界と、背景差分処理で生成したビデオアバタを

合成したものである。合成の際にはテクスチャマッピングを用いている。手順としては、まず背景差分処理において RGB 画像を RGBA 画像とする。A 値には差分結果が格納される。そして、A 値に対して閾値処理をし、0 または 1 のどちらかにする。次に閾値処理後の A 値と `glAlphaFunc()` を用いて描画する領域を決定する。その結果、ビデオアバタ画像の切り抜いた人物のみが描画され、もともと背景領域だった部分は透過することになる。

### 2.3 ビデオアバタ通信

生成したビデオアバタを使ってコミュニケーションをするための通信方式としては、TCP を使う場合と UDP を使う場合の 2 種類が考えられる。本研究では後者 UDP を用いて通信を行った。

UDP の特徴としては次の 2 つが挙げられる。1 つ目は、リアルタイム性があること。TCP はデータの到着を保障するため、ネットワークでパケットが消えると再送処理をする。この再送によって、リアルタイム性が損なわれる事がある。また、TCP はネットワークが混雑しているとデータ送信量を減らす輻輳制御を行う。この輻輳制御によっても TCP のリアルタイム性が損なわれている。これら再送や輻輳制御が UDP にはないため、UDP は TCP よりもリアルタイム性が高くなる。映像や音声などのデータを転送するには、すべてのデータが正しく転送されているよりも、最新のデータが転送されていることが重要である。2 つ目はデータが宛先に届いたかどうかを関知しないコレクションレス型のプロトコルであるため、TCP とは異なりデータの到着を保障しない。そのため、UDP ではネットワークでパケットが消えてしまった際の対処が必要になる。

上記 2 つの特徴のうち、本研究におけるビデオアバタは映像のリアルタイム通信であるから、前者が必要な条件となってくる。また、後者において本研究では 1 つ 1 つのパケットに順序をつけることで解決を図っている。1 フレームの映像データをいくつかの小さなデータに分割して送信をしているためである。順序付けをすることで、たとえ受信側ですべてのデータが正しく獲得できない場合でも、1 フレームという枠の中で正しい位置に分割データが格納される。そのため、全てのデータの受信が確認できなくても随時次のフレームへと更新を行うことが可能である。

## 3. CAVE 間ネットワーク

### 3.1 没入型ディスプレイ

没入型ディスプレイとはイリノイ大学の CAVE を代表とするスクリーンである。スクリーンを立方体状に配置しその中でユーザは立体視をする事で、臨場感の高い仮想空間を体験できる。本研究ではこの没入型ディスプレイをビデオアバタの投影環境として用いている。これによりビデオアバタで表現された人物像と 3 次元仮想世界の合成結果をより臨場感のある形で表現することが可能となる。

使用している没入型ディスプレイは、筑波大学の CS Gallery、東京大学の CABIN、そして京都大学の CAVE である。

### 3.2 JGN ネットワーク

ビデオアパタを用いて CAVE 間コミュニケーションを行おうとした場合、各地点にある CAVE 同士を広帯域のネットワークで結ぶ必要がある。本研究では、ネットワーク環境として JGN を使用した。これは JGN(Japan Gigabit Network: 研究開発用ギガビットネットワーク)を発展させた新たな超高速・高機能研究開発テストベッドネットワークとして、独立行政法人情報通信研究機構が運用しているオープンなテストベッドネットワーク環境である。全国規模の IP ネットワークで最大 20Gbps の回線を有し、高速大容量データを伝送交換できる環境を実現している。

### 3.3 接続方法

本研究では、つくば、東京、京都の 3 地点間をそれぞれ JGN で接続した。図 3 はそのネットワーク構成を示したものである。

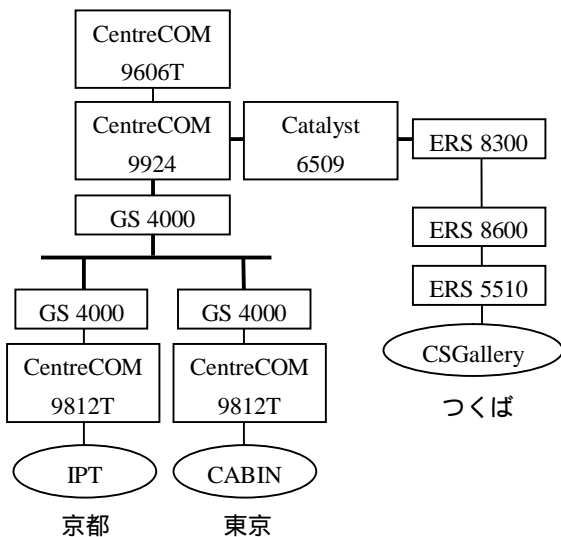


図 3 ネットワーク構成  
Fig.3 network composition.

つくばでは、JGN の拠点であるつくば JGN リサーチセンターから「つくば WAN」を経由して CS Gallery へとネットワークを延ばしている。ここで、「つくば WAN」とは筑波研究学園都市の研究機関を超高速(10G)のアクセスリングで結ぶネットワークである。東京では JGN のスイッチから直接 CABIN へと、そして京都では学内 LAN を経由して CAVE 型ディスプレイへとそれぞれネットワークを接続している。

このネットワーク環境を使用することで、つくば、東京、および京都の没入型ディスプレイ間が結ばれ、3 地点間でのデータ相互通信が可能となる。

## 4. 実験

### 4.1 実験環境

使用する没入型ディスプレイは、筑波大学の 3 面スクリーン構成の CS Gallery、東京大学の 5 面スクリーン構成の CABIN、そして京都大学の 3 面スクリーン構成の CAVE である。ただし、今回 CABIN は 5 面のうち正面、左面、そして床面スクリーンの 3 面のみを使用した。各地点それぞれには JGN ネットワークに接続している同機種の通信用 PC (Red Hat Enterprise Linux WS V4(EM64T) Intel Xeon, 3.0GHz) を配置している。その PC に CAVE 描画用 PC (つくば:3 台、東京:3 台、京都:3 台) が接続されている。それぞれの CAVE の描画はその台数に応じたクラスタ方式で行われている。

撮影環境としては SONY の DFW-X710 を使用している。この IEEE1394 カメラを JGN ネットワークにつながっている通信用 PC に取り付け、リアルタイムで撮影および通信を行う。上記ビデオカメラから得られる 1 フレームの画像は 640 × 480 ピクセルの RGB 画像である。また、毎秒 15 フレームの画像のデジタル出力をしている。

### 4.2 実験方法

処理の流れは以下の通りである。

通信用 PC はビデオカメラで撮影した映像を他の 2 地点へ送信する。

それぞれ他の 2 地点から送られてきた映像データを通信用 PC で受信する。

通信用 PC は受信した映像データを読み込んで、CAVE 描画用 PC へと転送する。

CAVE 描画用 PC で背景差分処理をして描画する。  
(東京のみ通信用 PC で背景差分処理)

### 4.3 結果

表 1 2 地点間映像転送の平均フレームレート値  
Table 1 Average frame rate of video transfer in 2 sites.

受信側	東京	京都
つくば	14.9	4.22

単位:fps

表 2 3 地点間映像転送の平均フレームレート値  
Table 2 Average frame rate of video transfer in 3 sites.

	つくば	東京	京都
つくば	- <sup>1)</sup>	3.28	3.29
東京	3.42	-	3.42
京都	3.21	3.21	-

単位:fps

表 3 CAVE 描画の平均フレームレート値  
Table 3 Average frame rate of CAVE drawing.

	つくば	東京	京都
正面	15.3	7.15	9.97
側面	15.3	7.15	9.92
床面	15.3	7.15	9.92

単位:fps

1: 「-」はデータ無し。他の地点からのみ受信をしている。

表1は2地点間の通信時においてつくばで描画をした際のフレームレートを示したものである。東京と京都の間にはフレームレートに大きな差が見られた。

表2、3はそれぞれ映像転送とCAVE内での描画のフレームレートを示したものである。今回の実験ではJGNで映像データを転送する部分と、受信した映像データをCAVE描画用PCへ転送する部分とを分けて実行した。その結果、各地点では映像データ転送の速度に依存しないでCAVEを描画することが可能となった。

表1より2地点間では、つくば-東大間で14.9fpsの通信が可能であった。しかし、表3からわかる通り3地点間にすることで3.42fpsまで低下してしまった。これは映像転送において撮影側が他の2地点へデータを送信する際に、その2地点両方からの影響を受けているためと考えられる。その結果、通信速度が速い地点であっても通信速度が遅い地点につられて、結果として両方が遅くなってしまいう状況ができてしまった。

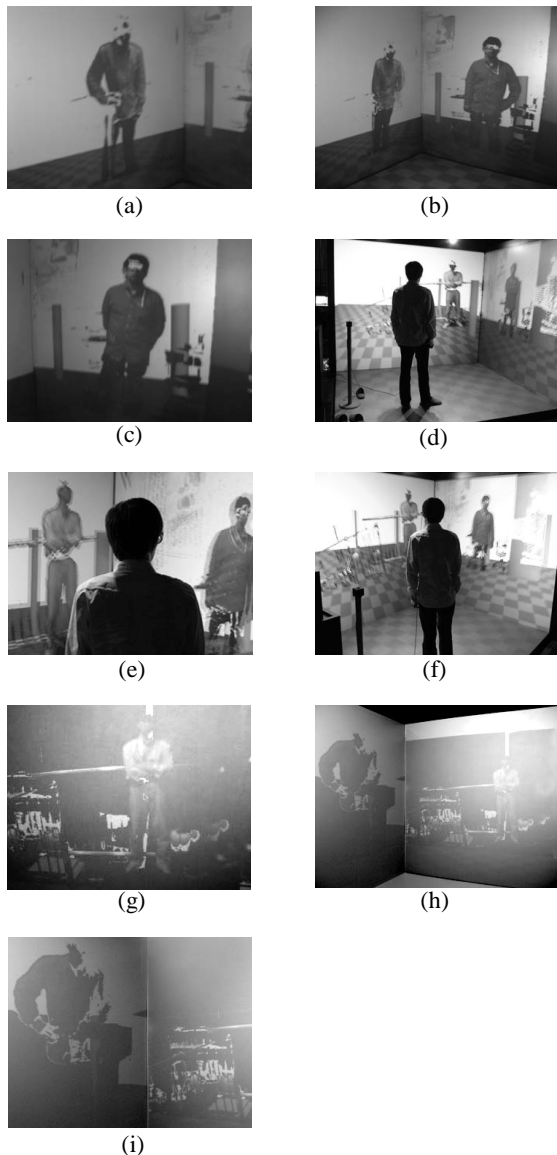


図4 CAVE 描画風景  
Fig.4 drawing of CAVE.

各 CAVE スクリーンの描画部分については安定した動作を確認することができた。図4が3地点にある各CAVEに描画した様子である。(a)~(c)が東大CABIN、(d)~(f)がつくばCS Gallery、(g)~(i)が京大CAVEの様子である。描画にはPCクラスタを用いていることから、各地点でCAVE内のデータ転送に関しても3台のPCを同期させて動作させる必要があった。表2よりその同期の問題に対してはサイト間の通信と独立に動作していることを確認することができた。

## 5. まとめ

本研究では、多地点間のCAVEを同時に動作してのビデオアパ通信を行った。特につくば、東京、そして京都の3つのCAVEをJANネットワークで接続することで、実時間でのビデオアパ通信を実現した。

今後の課題としては、ビデオアパを生成する際の背景差分法を改良し、鮮明な切抜きをすることがある。そして映像データ転送部分の速度の向上も必要である。より実時間に近い形での臨場感のあるコミュニケーション法を目指していく。

## 参考文献

- [1] 石井, 渡辺: Virtual Actorを対面合成した身体的ビデオコミュニケーションシステム; ヒューマンインタフェース学会論文誌, Vol.5, No.2, pp.73\_82 (2003)
- [2] 小木, 山田, 栗田, 服部, 廣瀬: 仮想空間共有のためのビデオアパ技術とその利用法; TVRSJ, Vol.8, No.1 (2003)
- [3] Cruz-Neir,C., Sandin,D.J., Defanti,T.A.: Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE; Proc. Of SIGGRAPH'93, pp.135\_142 (1993)
- [4] 小木: 没入型ディスプレイの特性と応用の展開; ヒューマンインタフェース学会論文誌, Vol.1, No.4, pp.43\_49 (1991)