

仮想空間共有のためのビデオアバタ技術とその利用法

小木哲朗^{*1,*2} 山田俊郎^{*1} 栗田裕二^{*1} 服部陽一^{*1} 廣瀬通孝^{*2}

Usage of Video Avatar Technology for Sharing Virtual World

Tetsuro Ogi^{*1,*2}, Toshio Yamada^{*1}, Yuji Kurita^{*1}, Yoichi Hattori^{*1}, Michitaka Hirose^{*2}

Abstract - Video avatar technology has become very popular as a high presence communication tool in the shared virtual world. In order to generate a realistic video avatar, several modeling techniques, such as the 2D plate model, 2.5D surface model or 3D voxel model have been proposed. This paper discusses the features and the usage of the various video avatar techniques. The appropriate video avatar that is used in the application system should be selected according to the purpose and the system environments. In addition, we also discuss the video avatar studio and the video avatar server technologies that were developed to utilize the various video avatar methods effectively in the networked virtual environments.

Keywords : video avatar, communication, shared virtual world, avatar server, avatar studio

1. はじめに

CAVEやCABINに代表される没入型ディスプレイの登場により、臨場感の高い仮想空間が提示されるようになってきた。そのため、仮想世界に合成される人物像に関しても、リアリティの高い表現手法が要求されている。一方、広帯域ネットワーク環境の整備に伴い、遠隔地間での仮想空間共有を目指した研究が盛んになってきた。そのため、ネットワーク上のコミュニケーション手段としても、リアリティの高い人物表現が望まれている。このような要求に応えるため、近年研究が盛んになってきた技術にビデオアバタがあげられる。ビデオアバタとは、ビデオ映像を用いることで3次元仮想空間内で臨場感の高い人物表現を行う方法である。図1は仮想世界に合成されたビデオアバタの例を示したものであるが、リアリティの高いアバタを生成するために、これまでに種々の方法が提案されている。

本論文では、これまでに提案されてきた各種のビデオアバタ技術の特徴について整理を行い、ビデオアバタの生成における技術課題の抽出とその利用方法について検討を行う。また、ここでの検討結果を踏まえ、ビデオアバタをより効果的に利用するために筆者等が現在開発を行っている、ビデオアバタスタジオ、ビデオアバタサーバの設計方針について論じる。



図1 没入型多面ディスプレイに提示されたビデオアバタ
Fig.1 Video Avatar Displayed in the Multi-screen Immersive Projection Display

2. ビデオアバタ技術の背景

ビデオアバタ技術の研究が盛んになってきた背景には、幾つかの流れが見られる。まず、共有仮想空間におけるコミュニケーション手法の研究分野があげられる。共有仮想空間のコミュニケーション手法としては、従来多くのシステムでCGを用いたアバタが用いられてきた。Habitat[1]等の初期のシステムではアバタ表現として2次元キャラクタが使われていたが、徐々に3次元モデルのアバタ[2][3]やリアリティの高いCGの人物表現[4]が使われるようになってきた。しかしポリゴンモデルだけでは表現能力に限界があり、顔画像のテクスチャ[5]や、顔のビデオ映像[6][7]を用いたシステムが現れてきた。このビデオ映像の利用をアバタの全身表現に拡張したものがビデオアバタと言える。

* 1: 通信・放送機構ぎふMVLリサーチセンター

* 2: 東京大学IML

* 1: Gifu MVL Research Center, TAO

* 2: IML, University of Tokyo

一方、ビデオ映像を用いたコミュニケーション手法としては、これまでもテレビ会議やCSCWの研究分野がある。従来のテレビ会議システムはお互いのビデオ映像を送受信するだけであったが、ここに空間共有という概念が導入されるようになってきた。例えば、ビデオ映像によるコミュニケーションを行いながら情報の共有を行ったり[8][9]、没入型多面ディスプレイを用いてお互いの位置関係を伴ったテレビ会議システムの研究[10]等が行われている。これらの空間共有の機能が增大してくると、コミュニケーションに使われるビデオ映像は、次第にアバタとしての性質を帯びてくるようになる。

また、コンピュータビジョンや画像処理の研究分野も、ビデオアバタ技術に大きな影響を与えている。例えば、これまでにステレオカメラによるデプス情報の取得[11]や、virtualized realityの研究[12]等、ビデオ映像に基づいた対象物の形状認識に関する研究が行われてきたが、これらの研究成果を人物の撮影に応用することで、ビデオアバタは2次元モデルから3次元モデルへと発展してきた。

3. ビデオアバタの技術課題

3.1 ビデオアバタの生成方法

ビデオアバタの生成に関してはこれまでに種々の方法が提案されているが、その基本的な処理は図2に示すことができる。まず利用者をビデオカメラで撮影し、この撮影映像から人物の形状モデルを作成すると同時に、人物映像の切り出しを行う。この形状モデルと人物映像が、ビデオアバタデータとしてネットワークを介して伝送、あるいはデータベースに格納される。仮想世界の生成サイトでは、これらのデータを受け取り、あるいはデータベースから読み込み、人物の形状モデルに切り出された人物映像をビデオテクスチャとしてテクスチャマッピングする等の方法で、ビデオアバタを生成する。特に、遠

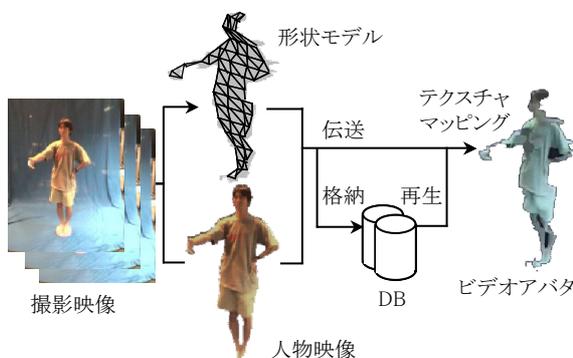


図2 ビデオアバタの生成方法
Fig.2 Creation Method of Video Avatar

隔地間の両サイトでこれらの処理を相互に実行することで、ビデオアバタをコミュニケーション手段として利用することができる。

臨場感の高いビデオアバタを生成するためには、上記の生成過程において、いかに精巧な人物の形状モデルを作成するか、あるいはいかにきれいな人物の切り出し映像を作成するかが技術的な課題となる。そのため、これまでに提案されてきた種々のビデオアバタ生成手法は、形状モデルの作成と人物映像の作成という、2つの技術的な視点から整理を行うことができる。

3.2 形状モデルの作成

ビデオカメラで撮影される映像は2次元情報であるが、人間の身体は3次元形状を有している。そのため、ビデオアバタの生成においては、いかに3次元情報を持つ形状モデルを作成できるかが重要となる。

(1) 2次元モデル

最も単純な形状モデルは2次元の平板モデルを用いる方法である[13][14]。この方法では、撮影されたビデオ映像がそのまま短冊状の平板の上にテクスチャマッピングされる。そのため、平板上の2次元的な人物の動作は表現できるが、前方に手を伸ばす等の奥行き方向に関する動作を表現することはできない。

2次元平板モデルを用いながら3次元情報を付加する方法としては、カメラ切り替えを用いる方法があげられる。これは図3に示すように、多数の視点方向から人物像を撮影し、相手の視点移動に追従して最も近い視点方向からの撮影映像に切り替える方法である[15][16]。1台のカメラから得られる映像は2次元の平面情報であるが、視点の移動に応じてカメラを切り替えることで、運動視差の効果を表わすことができる。映画の特殊撮影で用いられるbullet-time shotはこの手法の一つの応用と考えられるが[17]、滑らかな視点移動を表現するためには、人物を取り囲むように多数のカメラを配置する必要がある。

また2次元モデルを用いながら、ステレオカメラを用いることで視差情報を付加する方法もある

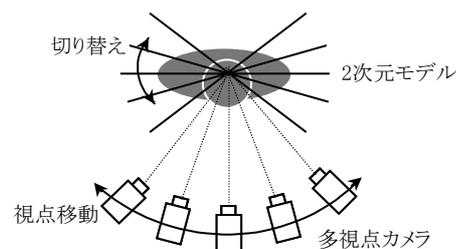


図3 2次元モデルによるカメラ切り替え
Fig.3 Camera Switching Using 2D Avatar Model

[18][19]. ステレオカメラで撮影された視差映像を立体視ディスプレイに提示することで、利用者はビデオアバターの立体映像を見ることができる。しかしながらこの方法は、撮影カメラの視点位置でしか正しい立体視映像を見ることができないため、利用者がディスプレイ空間内を大きく移動するような場合には不向きである。

(2) 2.5次元モデル

ビデオ映像から両眼視差と運動視差の両方を備えたモデルを生成する方法としては、ステレオマッチングのアルゴリズムを用いる方法がある[20][21]。この方法では、ステレオカメラで撮影された視差映像から撮影対象の各ピクセルに対する距離情報を算出する。この際、計測精度を上げるために明滅光のパターンを照射する等の改良法もある[22]。この距離情報から人物の表面形状モデルを作成し、その上にビデオ映像をテクスチャマッピングすることで、立体形状を持つビデオアバターを生成することができる。しかし、この方法ではカメラの撮影面に対してしか形状モデルを生成することができないため、通称2.5次元ビデオアバターと呼ばれている。

2.5次元ビデオアバターはカメラの撮影方向から見る場合には自然な立体映像に見えるが、利用者が相手の側面や後ろ側に移動する場合には、形状モデルを持たないため張りぼての人物像になってしまう。そのため、複数台のステレオカメラで人物像を撮影し、各方向から作成した2.5次元モデルを利用者の視点移動に応じて切り替える方法(図4)[23]や、マルチカメラによる撮影映像から2台ずつをステレオカメラのペアとして使用し、全体として広範囲の表面形状モデルを作成する方法[22]等が提案されている。

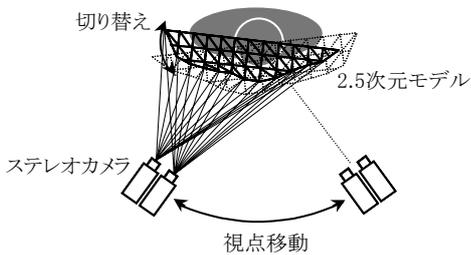


図4 ステレオカメラ切り替えによる2.5次元アバターモデル
Fig.4 Stereo Camera Switching Using 2.5D Avatar Model

(3) 3次元モデル

2.5次元モデルの不完全さを解決するために、ビデオ映像から完全な3次元モデルを生成する方法も提案されている。例えば、ステレオマッチングの手法を3次元に拡張した方法として、複数台のステレオカ

メラによる表面形状モデルを統合する方法があげられる[24]。この方法は、カメラ間の正確なキャリブレーションが必要になるが、全体として全周形状を持つポリゴンモデルを生成することができる。

また、多数のカメラによる撮影映像から安定な3次元モデル生成する手法として、マルチベースラインステレオマッチングと呼ばれる方法も提案されている[25]。この方法は、複数台のカメラ映像から対応点抽出を行い、各カメラで得られる距離情報を空間的に融合することで、全体として誤差の少ない形状モデルを生成することができるが、リアルタイムで処理を行うには計算負荷が大きいという問題がある。

これに対してマルチカメラの映像からボクセルモデルを作成する視体積交差法と呼ばれる方法も提案されている[26][27]。これは撮影対象となる空間を微小なボクセルに分割し、各ボクセルがそれぞれのカメラで撮影されるシルエットの内側か外側かを判定することで、3次元のボクセルモデルを構築する手法である(図5)。この方法は比較的計算負荷が小さく、ボクセルからポリゴンを生成することで、ビデオアバターのモデルとして利用することができる。

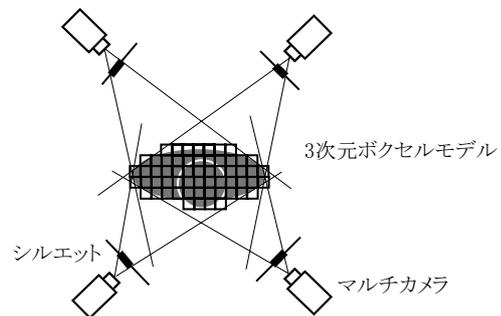


図5 マルチカメラシステムによる3次元ボクセルモデル
Fig.5 3D Voxel Model Made by Multi-camera System

また同様に視体積交差法を用いながらポリゴンを生成せずに、任意視点映像に対する各ピクセル値をボクセルデータから直接求める方法も提案されているが[28]、立体視に対応するためには視差ごとに映像を生成しなければならない等の問題がある。

上述のような撮影映像から人物の形状モデルをリアルタイムで作成する方法は、モデルの精度が撮影条件に大きく依存する。そのため、あらかじめ顔等の形状モデルをオフラインで作成しておく方法も用いられている。これは、Talking Heads[29]に見られる手法の応用と考えられるが、高精細なアバターを生成するために、リアルタイムで更新を行うビデオテクスチャを目や口の部分に限定したり[30]、視点方向によるカメラ映像の切り替え[31]、複数カメラの映像の重ね合わせ[32]、等の各種の工夫が行われている。こ

表1 ビデオアバタの各種形状モデルの特徴
Table 1 Features of Several Video Avatar Models

	アバタモデル	両眼視差	視点移動	画質	全身表現	視野遮蔽	計算負荷	コスト
2D	平板モデル	×	×	○	○	○	○	○
	平板+切替	×	○	○	○	×	○	○
	平板ステレオ視	○	×	○	○	○	○	○
2.5D	表面形状モデル	○	△	△	○	△	○	△
	表面形状+切替	○	○	△	○	△	○	△
	広範囲表面形状	○	○	△	○	×	○	△
3D	表面形状+融合	○	○	△	○	△	△	△
	マルチベースライン	○	○	△	○	×	×	×
	視体積交差+ポリゴン	○	○	△	○	×	△	×
	視体積交差+ボクセル	△	○	△	○	×	△	×
	顔モデル	○	△	○	×	○	○	○
	顔+切替	○	○	○	×	△	○	○
	顔+融合	○	○	○	×	△	○	○

これらの方法は、モデルとテクスチャの正確な位置合わせを行うことが必要になるが、顔等の比較的形狀変化の少ない部分には有効な方法と言える。

表1は、以上の各種形状モデルに関する特徴を整理したものである。ここでは、各項目について3段階の相対的な評価を記した。この表から、それぞれのモデルには特徴があり、利用目的と環境に応じて適当なモデルを選択する必要があることが分かる。ビデオアバタの応用と使用すべきモデルに関しては、4章で改めて議論を行うことにする。

3.2 人物映像の作成

ビデオアバタの技術要素としては、次にテクスチャとなる人物映像の作成が必要であるが、ここではできるだけきれいな映像を得るために、人物像の切り抜き方法、撮影方法等が問題になっている。

(1) 人物像の切り抜き

撮影映像から人物像だけを切り抜く方法は、撮影環境によって種々の手法が使用されている。最も代表的な手法はクロマキーを用いる方法である[13][14][19]。この方法は、切り出される人物像のクオリティは高いが、背景にブルーバックを使用できる特殊な撮影環境を必要とするという制約がある。

背景にブルーバックを使用できない場合の代表的な方法としては、背景差分法がある[27][31]。これは、あらかじめ人物のいない背景画像を撮影しておき、人物の撮影映像との差分を取ることで人物像を切り出す方法である。この方法は、光源の移動等による背景映像の変化に影響を受けるため、正規化色空間で差分を取る等の改良法も提案されている[21]。背景差分法は、比較的安定な切り抜きを行うことができるが、全天周型ディスプレイ等で背面にもスクリー

ンが置かれる場合には、背景映像が変化するため使用することができない。

また、背景が変化しても切り抜きが可能な方法として、デプスキーを用いる方法があげられる[11][22]。これはカメラから撮影対象までの距離（デプス）情報を利用して、特定の対象物を切り抜く方法である。ステレオカメラを用いる場合には、ステレオマッチングのアルゴリズムにより各ピクセルごとの距離を計測することができるため、これをデプスキーとして利用することができる。この方法は、背景が変化する場合でも利用することができるが、切り抜きの画像品質はデプスの精度にそのまま依存するため、照明等の撮影条件を十分に調整する必要がある。

また、同様に全天周型のディスプレイ環境で利用可能な方法に熱画像を用いる方法がある。これはサーマルカメラを用いて対象物の表面温度を計測することで、温度をキーにして人物を切り抜く方法である[33]。この方法は、背景の変化や照明条件等に影響を受けにくい、サーマルカメラという特殊なカメラを用意することが必要になる。

(2) 撮影方法

きれいな人物映像を作成するためには、切り抜き方法だけではなく、映像そのものを鮮明に撮影しなければならない。特に、没入型ディスプレイ環境の利用者を撮影するためには、広画角、高感度のカメラを使用することが必要となり、より効果的な撮影を行うために、カメラ配置や撮影手法に関する種々の提案が行われている。

例えば、人物がある空間内を移動する場合、カメラはその行動範囲全体で人物映像を撮影できなければならない。そのため、人物の位置をセンサで計測し、それに追従するようにカメラの首振り制御す

る方法等が提案されている[34][35].

また、プロジェクタを用いた仮想環境で撮影を行う場合には、照明の問題も重要である。映像を提示するためにはディスプレイ空間内は暗い方が好ましいが、人物像を撮影するためには、ある程度の明るさが必要である。そのため、液晶シャッターメガネの開閉に同期させて、ストロボを照射させる同期式の照明方法等が提案されている。具体的なシステムとしては、液晶シャッターメガネ、ストロボ照明に合わせてクロマキー背景、カメラ撮影を同期させる方法[36]、更にスイッチ式スクリーンを同期させスクリーンの背後からカメラ撮影を行う方法[37]等が提案されている。この同期式照明方法は、周りの観察者に対する明かりのちらつき等の影響はあるが、利用者に対しては効果的な照明を与えることができる。

このように、人物像の切り抜き方法や撮影方法は、撮影スタジオやディスプレイ環境によって利用できる手法が制限され、適する手法も変わってくる。そのため、あるアプリケーションで利用すべきビデオアバタ手法は、その目的やシステム環境に合わせて最適な方法を選択する必要があると言える。

4. ビデオアバタの応用分野

4.1 対称型コミュニケーション

ビデオアバタの応用分野は、その発展経緯から種々の領域で応用が試みられている。ここでは、各種の応用分野とそこで使用すべき、適したビデオアバタ手法の関係について考察を行う。

まずビデオアバタの代表的な応用分野として、共有空間におけるコミュニケーションがあげられる。例えば、遠隔地間での空間共有型のオフィス[38][39]や会議システムへの応用である。この種のシステムでは、遠隔地の利用者が対等な関係でコミュニケーションを行い、アバタの撮影システムをディスプレイ内に設置する必要があるため、視界を妨げる大掛かりな装置は使用できない。また利用者は移動が少なく、主に対面でコミュニケーションを行うという特徴がある。そのため、ビデオアバタとしてはお互いの表情を表現でき、できるだけ画質の高い手法を用いることが望ましい。

また、共有仮想空間内に必要な情報を合成することで、設計、解析、可視化[40]等の各種の協調作業に応用することができる。これらの応用システムでは、アバタによる指差し等の空間的な情報伝達や動作表現を行えることが要求される。そのため2.5次元モデル等の立体視情報を有効に利用できるビデオアバタ手法を用いることが望ましいと考えられる。

一方、仮想都市等のサイバースペース上での多人数のコミュニケーションにビデオアバタを用いる方法も考えられている[6]。この場合、双方向コミュニケーションではあるが、共有空間には多数の利用者が参加するため、1対1の対話システムのようにお互いに正面を向いたコミュニケーションだけではなく、いろいろな方向から人物像を眺められることが必要となる。そのため、カメラ切り替え等の手法により多方向からの視点移動に対応できるモデルを用いることが望ましいと言える。

4.2 非対称型コミュニケーション

空間共有型のアプリケーションへの応用としては、遠隔プレゼンテーション[41]、遠隔教育[42]、遠隔操作指示等の応用分野への適用も行われている。これらのシステムでは、両者のコミュニケーションは対等な関係ではなく、教示者から被教示者への情報伝達が優先される非対称型のコミュニケーションとなる。そのため、教示者側から送信されるビデオアバタに対しては、マルチカメラシステムやブルーバックを使用する等、できるだけ映像品質の高い手法を用いることが必要であるが、被教示者側からのビデオアバタには、それ程高い品質は要求されない。

インタラクティブアート等の分野において、アーティストと鑑賞者のインタラクションが行われる場合も非対称型のコミュニケーションと言える[28][43]。この場合、アーティスト側から送られる映像は必ずしもリアリティが優先される訳ではないが、作品として要求されるビデオアバタの品質が優先的に確保されることが必要となる。

4.3 記録再生型コミュニケーション

また、利用者とアバタがリアルタイムの対話を必要とせず、一方向的な情報伝達だけが行われる場合もある。例えば、仮想博物館におけるギャラリートーク[44]等、仮想世界のナビゲーションにビデオアバタを利用する場合である。これらの応用分野では、リアルタイムでビデオアバタを伝送する必要はないため、あらかじめ記録しておいたアバタデータを再生する方法を用いることができる。そのため、ビデオアバタの生成手法としては、マルチカメラを備えたスタジオ等を使用し、高品質のビデオアバタ手法を用いることができる。

また双方向コミュニケーション環境において、サイト間で時差がある場合等に、アバタデータを一旦記録しておき必要な時に再生するという時間差を利用したコミュニケーション方法の提案も行われている[45]。この場合は記録再生型ではあるが、各利用者

のアバタデータは、それぞれ仮想世界の体験時に記録されるため、マルチカメラ等のディスプレイ映像を妨げるような手法は用いることができない。

5. MVLプロジェクトにおけるビデオアバタ技術

これまでの技術的な課題と応用に関する検討から、ビデオアバタは利用目的と環境に応じて適する手法を選択できることが望ましいと言える。ここでは、筆者等が現在行っているMVL (Multimedia Virtual Laboratory) プロジェクトを例に、各種のビデオアバタ技術をどのように利用すべきかという具体的な例を示す。

5.1 MVL研究環境

MVLとは、総務省が進めている広帯域ネットワーク上に展開される仮想の研究所の概念であるが、ここでは遠隔地の研究者が臨場感の高いコミュニケーションを行えることが求められる。このような機能を実現するため、筆者等は東京大学のCABIN[46]と岐阜県テクノプラザのCOSMOS[47]の没入型ディスプレイ間、あるいは多面型ワークベンチ[48]や実空間の作業室等を通信放送機構のJGN (Japan Gigabit Network) で接続した共有空間環境を構築した。このネットワーク上で高臨場感コミュニケーションを実現するための手段として、これまでに各種のビデオアバタ技術の開発を行ってきた。

5.2 MVLのためのビデオアバタ手法

MVLで利用するビデオアバタ手法としては、まず2次元モデルのカメラ切り替え手法を構築した。ここではブルーバックのスタジオを使用し、人物の周りに配置したマルチカメラで人物像を10度間隔で撮影し記録した。仮想空間の中では相手の視点位置に応じて最も近いフレームが選択され、仮想世界に合成される。この方法は、2次元モデルではあるが、視点の移動に対応し映像的にきれいな人物表現をできるため、各種のデモンストレーションシステムに埋め込み、仮想世界の案内役等の用途として使用している (図6)。

次にステレオカメラを用いた2.5次元ビデオアバタ手法の開発を行った。ここでは、3眼のステレオカメラであるPoint Grey Research社のTriclopsを使用して作成されたビデオアバタを、CABIN-COSMOS間のコミュニケーションに利用している[20]。2.5次元ビデオアバタはカメラ方向の表面形状モデルしか持たないが、指差し動作等の位置情報伝達には有効である。そのため、協調設計、データ可視化等、共有仮想空間内での協調作業への応用に有効な方法として

使用している (図7)。



図6 カメラ切り替えによる2次元モデルビデオアバタ
Fig.6 2D Video Avatar Using Camera Switching Method.



図7 2.5次元ビデオアバタを用いた可視化の協調作業
Fig.7 Collaborative Visualization Using 2.5D Video Avatar

また実空間の符号化伝送を目的としたビデオアバタ技術として、マルチカメラによる3次元モデル手法の開発を行っている[27]。この方法では、作業室内の人物の行動をマルチカメラで撮影し、視体積交差法からポリゴンモデルを構築する方法を用いている。マルチカメラが設置された特別な部屋を使用するため、双方向のコミュニケーションには利用できないが、実空間での人間の行動を3次元情報として記録したり、実空間の出来事を伝送して仮想空間に合成する等の目的に利用している (図8)。

また、顔の表情を伝達可能なビデオアバタとして、高解像度顔アバタ手法の開発を行った[32]。この方法では、あらかじめモデリングされた顔の形状モデルに、Optotrakでトラッキングされた利用者の顔を2方向から撮影したビデオ映像を用い、テクスチャマッピングを行っている (図9)。顔アバタは、ワークベンチでの作業等、利用者が大きく移動せずに、比較的近距离で相手とコミュニケーションを行う場合に有効な方法として利用している。



図8 3次元ビデオアバター手法を用いた実空間の伝送
Fig.8 Transmission of Real World Using 3D Avatar Technique



図11 18眼カメラを備えたビデオアバタスタジオ
Fig.11 Video Avatar Studio Equipped with 18 Cameras



図9 高精細3次元顔アバタを用いた対話
Fig.9 Communication Using High-resolution Facial Avatar

5.3 ビデオアバタスタジオ

MVLでは、上述した各種ビデオアバター手法をそれぞれの用途で使用することを目指しているが、これらに用途に応じて使い分けるためには、種々のビデオアバタを生成可能な撮影伝送システムが必要である。筆者等は、このような目的で現在ビデオアバタ生成用のスタジオの開発を行っている[49]。図10は、ビデオアバタスタジオの構成、図11はビデオアバタスタジオの様子を示したものである。

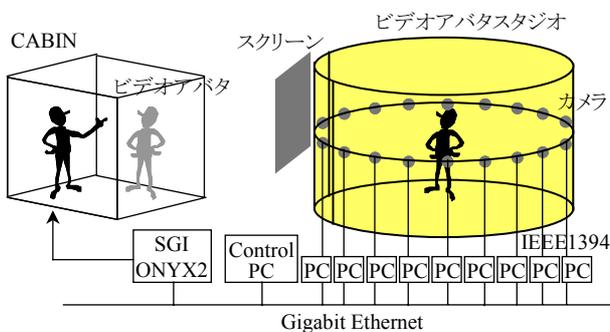


図10 ビデオアバタスタジオの構成
Fig.10 System Construction of Video Avatar Studio

このスタジオは、直径400cm、高さ257cmの円筒形の部屋で、ブルーバックの壁面に20度間隔で合計18台のCCDカメラ（SONY DFW-X700）が埋め込まれている。カメラの位置は撮影対象に合わせて、床から60cm、120cm、157.5cmの3通りの位置に付け換えて使用することができる。カメラで撮影された映像は、それぞれIEEE 1394接続でPCに送られ、ここでビデオアバタデータの作成が行われる。また、正面の壁面パネルは取り外しが可能になっており、これを外すと正面には100インチのスクリーンが設置されている。記録型のビデオアバタデータを作成する場合、あるいはスタジオから一方向的にビデオアバタの送信を行う場合には、壁面を閉じて全周方向の撮影映像を用いるが、共有仮想空間でのコミュニケーションにビデオアバタを用いる場合には、相手の映像を提示するために正面のスクリーンを使用する。

特定のビデオアバタを生成するための撮影スタジオはこれまでも幾つか作られているが[50][51]、本システムは複数台のカメラ映像に対して記録あるいは実時間画像処理を行うことで、種々のビデオアバタの生成に対応できることが特徴である。例えば、2次元モデルを用いる場合には、利用者の視点移動に応じて18方向からの映像切り替えを行う。また、隣り合うカメラをステレオカメラとして利用することで、奥行き情報を持った2.5次元ビデオアバタを生成することができる。また、全周の撮影映像を同時に使用することで、ボクセルモデルを用いた3次元ビデオアバタを生成することも可能になる。

5.4 ビデオアバタサーバ

次に、ビデオアバタデータの配信用として、ビデオアバタサーバの開発を行っている。これまでの多くのビデオアバタは、同種のシステムを持つ2地点間のコミュニケーションに利用されてきたが、このビデオアバタサーバは、多地点間でのコミュニケーシ

ョンを可能にするるとともに、各サイト間での異なるアバタモデルの使用、コミュニケーション型と記録再生型のビデオアバタの混在等、多様なビデオアバタの利用に対応することを目指している[52].

ネットワーク仮想空間の接続法やサーバの構築法に関してはこれまでも種々の研究が行われているが[53][54], 本システムは効率的なデータ通信を行いながら多様なアバタ手法を利用可能にすることが特徴である. 図12は現在開発中のビデオアバタサーバの機能構成を図示したものである.

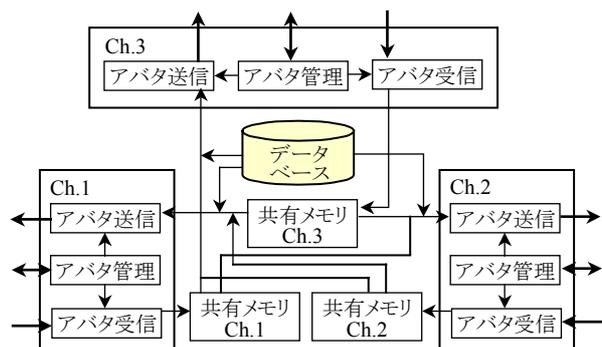


図12 ビデオアバタサーバの構成

Fig.12 Software Construction of Video Avatar Server

ビデオアバタサーバを用いた通信では、まず各サイトがサーバに接続要求を行う。サーバ側では、アバタ管理用プロセス、アバタ受信用プロセス、アバタ送信用プロセスが起動され、各クライアントと使用するアバタデータの情報や既に接続されている他のクライアント情報等を交換する。各サイトではアバタデータを一旦サーバへ送信し、サーバは接続されている各クライアントに必要なアバタデータを配送する。また、ビデオアバタサーバはリアルタイムのコミュニケーションをサポートするだけでなく、あらかじめ記録されたビデオアバタの再生にも対応する。この場合、アバタ送信プロセスはアバタデータをデータベースから読み込み、実時間速度での再生制御を行いながら配送することで、クライアント側では伝送データか記録データかの区別をすることなくビデオアバタの合成を行うことが可能になる。

以上のようなビデオアバタスタジオ、ビデオアバタサーバの技術を実現することで、ネットワーク上の各サイトではシステム環境と利用目的に応じ、より効果的なビデオアバタ手法を選択し使用することが可能になると期待される。

6. おわりに

本論文では、近年研究が盛んになってきたビデオ

アバタ技術について、各種生成手法の特徴、問題点、応用分野等の整理を行った。ビデオアバタの表現手法に関しては、2次元表現から2.5次元表現、3次元表現と表現能力が拡大しているが、特殊なカメラ装置の使用等、コストや利用性の問題があるため、現状では利用目的とシステム環境に応じた生成手法を選択することが望ましいと言える。またそのためには、各種のビデオアバタ手法の選択や混在が可能な利用環境を構築しておくことが必要である。

本論文の後半では、このような視点に立って現在開発を行っている、ビデオアバタスタジオとビデオアバタサーバについて説明を行った。これらのシステムについては、現在実装中の機能もあるため、ここでは設計方針を述べるにとどめたが、これらの有効性評価等については改めて報告を行いたい。

参考文献

- [1] 福田, 田原: ビジュアルパソコン通信富士通Habitatにおけるオンライン処理技術, FUJITSU, Vol.43, No.6, pp.639-646 (1992).
- [2] 松田, 三宅: パーソナルエージェント指向仮想社会PAW (第2版) の構築と評価, 情報処理学会論文誌, Vol.41, No.10, pp.2698-2707 (2000).
- [3] Leigh, J., Johnson, A.: Supporting Transcontinental Collaborative Work in Persistent Virtual Environments, IEEE Computer Graphics and Applications, Vol.16, No.4, pp.47-51 (1996).
- [4] Kouno, T., Honda, S., Suzuki, Y., Ishibashi, S.: Immersion-type Shared Communication Environment, IEEE PCM 2000, pp.102-105 (2000)
- [5] Morishima, S., Yotsukura, T.: Face-to-face Communicative Avatar Driven by Voice, IEEE ICIP'99 (1999).
- [6] Sugawara, S., Suzuki, G., Nagashima, Y., Matsuura, M., Tanigawa, H., Moriuchi, M.: InterSpace- Networked Virtual World for Visual Communication, IEICE Trans. Information & Systems, Vol.E-77D, No.12, pp.1344-1349 (1994).
- [7] 中西, 吉田, 西村, 石田: FreeWalk: 3次元仮想空間を用いた非形式的なコミュニケーションの支援, 情報処理学会論文誌, Vol.39, No.5, pp.1356-1364 (1998).
- [8] Kobayashi M., Ishii H., ClearBoard: A Novel Shared Drawing Medium that Supports Gaze Awareness in Remote Collaboration, IEICE Trans. Communications, Vol.E76-B, No.6, pp.609-617 (1993).
- [9] Yamashita, J., Kuzuoka, H., Yamazaki, K., Miki, H., Yamazaki, A., Kato, H., Suzuki, H.: Agora: Supporting Multi-participant Telecollaboration, HCI International 99, Vol.2, pp.543-547 (1999).
- [10] 鈴木, 鈴木, 田中, 荒川, 小木: 4者間討議空間システムの構築, HI学会研究報告集, Vol.1, No.2, pp.39-42 (1999).
- [11] 桑島: 世界初のPCベース・フルカラーステレオ・ビジョン・システムCOLOR TRICLOPS, 映像情報インダストリアル, Vol.31, No.4, pp.17-22 (1999).
- [12] Kanade, T., Rander P.: Virtualized Reality: Being Mobile in a Visual Scene, ICAT/VRST'95, pp.133-142 (1995).

- [13] Insley J., Sandin D., DeFanti, T.: Using Video to Create Avatars in Virtual Reality, Visual Proceedings of 1997 SIGGRAPH, pp.128 (1997).
- [14] Keum, S., Park, J., Won, Y., Park Y.: Real-Image-Based Distributed Virtual Reality System with Java3D, ICAT2001 (2001).
- [15] 小木, 山田, 玉川, 廣瀬: 共有没入空間におけるビデオアバタを用いた位置関係表現, 電子情報通信学会論文誌, Vol.J82-D-II, No.10, pp.1826-1834 (1999).
- [16] 鈴木: 3D Replayのテレビ放映, 日本VR学会誌, Vol.7, No.2, pp.56-57 (2002).
- [17] 大口: 「マトリックス」のVFX (視覚効果) について, <http://japan.whatisthematrix.com/vfx.html>
- [18] 鈴木, 田中, 川嶋, 荒川: NetUNIVERSを用いた3D空間共有通信システム, 2001年電子情報通信学会総合大会, A-16-35 (2001).
- [19] Hasenbrink, F., Lalioti, V.: Towards Immersive Telepresence SCHLOSSTAG'97, IPT98 (1998).
- [20] Tamagawa, K., Yamada, T., Ogi, T., Hirose, M.: Developing a 2.5-D Video Avatar, IEEE Signal Processing Magazine, Vol.18, No.3, pp.35-42 (2001).
- [21] Kim, N., Woo, W., Tadenuma, M.: Photo-realistic Interactive Virtual Environment Generation Using Multiview Cameras, SPIE PW-EI-VCIP'01, Vol.4310, pp.245-254 (2001).
- [22] Lanier, J.: Virtually There, Scientific American, April, pp.66-75 (2001).
- [23] Ogi, T., Yamada, T., Tamagawa, K., Kano, M., Hirose, M.: Immersive Telecommunication Using Stereo Video Avatar, IEEE VR2001, pp.45-51 (2001).
- [24] Sein, M.M., Suzuki, Y., Kakeya, H., Arakawa, Y.: Generating the 3D Model of an Object for Realizing the 3D Shape-shared Communication over the Network, IEEE KMN2002 (2002).
- [25] Vedula, S., Rander, P., Saito, H., Kanade, T.: Modeling, Combining, and Rendering Dynamic Real-World Events From Image Sequences, VSMM98, pp.326-332 (1998).
- [26] Moezzi, S., Katkere, A., Kuramura, D.Y., Jain, R.: Immersive Video, VRAIS'96, pp.17-24 (1996).
- [27] 栗田, 國枝: 多視点カメラによる3次元モデル生成とネットワーク伝送システム, 日本VR学会第7回大会論文集, pp.489-492 (2002).
- [28] Prince, S., Cheok A.D., Williamson T., Johnson N., Kato H., Farbiz, F.: 3D Live: Real Time Captured Content for Mixed Reality, ISMAR2002 (2002).
- [29] Negroponte, N., Parker, W.: Talking Heads: Display Technology for Persona, Society for Information Display International Seminar-Symposium-Exhibition (1981).
- [30] Leung, W., Tseng, B., Shae, Z., Hendriks, F., Chen, T.: Realistic Video Avatar, IEEE ICME2000, Vol.II, pp.631-634 (2000).
- [31] Rajan, V., Subramanian, S., Keenan, D., Johnson A., Sandin, D., DeFanti, T.: A Realistic Video Avatar System for Networked Virtual Environments, IPT2002 (2002).
- [32] 井原, 服部: 没入型ディスプレイ環境を想定した写実的顔アバタのリアルタイム制御, 3次元画像コンファレンス2001講演論文集, pp.25-28 (2001).
- [33] 川原, 松下, 新田, 苗村, 原島: 透過型ビデオアバタの提案と実装~熱画像を利用した実時間システム~, 日本VR学会第5回大会論文集, pp.333-336 (2000).
- [34] 中北, 町田, 竹村, 横矢: 首振りカメラを用いた人物位置に追従するビデオアバタ表示, 信学技法, EID, Vol.100, No.605, pp.31-36 (2000).
- [35] 広瀬, 山田, 小木, 廣瀬: カメラ追従による広視野ビデオアバタ手法, HI学会研究報告集, Vol.3, No.2, pp.33-36 (2001).
- [36] 廣瀬, 小木, 加納, 山田: 没入型ディスプレイ間通信のための同期式クロマキー手法, HI学会研究報告集, Vol.2, No.2, pp.49-52 (2000).
- [37] Kunz, A., Spagno, C.P.: Simultaneous Projection and Picture Acquisition for a Distributed Collaborative Environment, IEEE VR 2002, pp.279-280 (2002).
- [38] Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., Fuchs, H.: The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays, SIGGRAPH98, pp.179-188 (1998).
- [39] Briteneder, C., Gibbs, S., Arapis, C.: TELEPORT- An Augmented Reality Teleconferencing Environment, 3rd Eurographics Workshop on Virtual Environments Coexistence & Collaboration (1996).
- [40] Ogi, T., Tamagawa, K., Hirose, M.: Collaborative Scientific Visualization in Networked Immersive Virtual Environment, SMC'99, (1999).
- [41] 小木, 山田, 廣瀬: 共有没入仮想空間を用いた遠隔プレゼンテーション, HIS2002, pp.347-350 (2002).
- [42] Yang, X., Goebbels, G.: Distributed Medical Teaching through Responsive Workbench and Cyberstage, ICIG2000 (2000).
- [43] Macintyre, B., Lohse, M., Bolter, J., Moreno, E.: Ghosts in the Machine: Integrating 2D Video Actors into a 3D AR System, ISMR 2001, pp.73-80 (2001).
- [44] Yura, S., Usaka, T., Sakamura, K.: Video Avatar: Embedded Video for Collaborative Virtual Environment, IEEE ICMCS, Vol.2, pp.433-438 (1999).
- [45] Imai, T., Johnson, A., Leigh, J., Pape, D., DeFanti, T.: The Virtual Mail System, IEEE VR'99, pp.78 (1999).
- [46] 廣瀬, 小木, 石綿, 山田: 多面型全天周ディスプレイ (CABIN) の開発とその特性評価, 電子情報通信学会論文誌D-II, Vol.J81-D-II, No.5, pp.888-896, (1998).
- [47] 山田, 棚橋, 小木, 廣瀬: 完全没入型6面ディスプレイCOSMOSの開発と空間ナビゲーションにおける効果, 日本VR学会論文誌, Vol.4, No.3, pp.531-538 (1999).
- [48] Yamada, T., Tsubouchi, D., Ogi, T., Hirose, M.: Desk-sized Immersive Workplace Using Force Feedback Grid Interface, IEEE VR 2002, pp.135-142 (2002).
- [49] 広瀬, 廣瀬, 広田, 小木: 多視点映像から円筒座標系への投影による3次元形状生成, 日本VR学会第7回大会論文集, pp.499-502 (2002).
- [50] Kanade, T., Saito, H., Vedula, S.: The 3D Room: Digitizing Time-Varying 3D Events by Synchronized Multiple Video Streams, Robotics Institute Technical Report, CMU-RI-TR-98-34 (1998).
- [51] Utsugi, K., Beniyama, F., Namai, H., Moriya, T., Takeda, H.: A High-Resolution Avatar System Using Partial Compositions, ICAT2002, pp.65-71 (2002).
- [52] 山田, 小木, 廣瀬: ビデオアバタコミュニケーションサーバの開発, 日本VR学会7回大会論文集, pp.493-494 (2002).
- [53] 清川, 竹村: ネットワークバーチャルリアリティシステムの構成法, 情報処理, Vol.42, No.3, pp.251-256 (2001).
- [54] 清末, 湯田, 山名, 加藤, 正木, 一之瀬: クライアントの性能とサービスの多様性に対応した3次元サイバースペースシステムの機能分散型サーバアーキテクチャの提案, 日本バーチャルリアリティ学会論文誌, Vol.4, No.2, pp.351-356 (1999).

(2002年11月21日受付)

[著者紹介]

小木哲朗

(正会員)



1986年東京大学大学院工学系研究科修士課程修了. 同年(株)三菱総合研究所入社. 1994年東京大学大学院工学系研究科博士課程修了. 1996年東京大学IML助教授. 1999年より通信・放送機構研究員, 東京大学IML客員研究員. ビジュアルライゼーション, 臨場感通信の研究に従事. 博士(工学).

山田俊郎

(正会員)



1994年福井大学大学院工学研究科博士前期課程修了. 現在, 岐阜県生産情報技術研究所, 通信・放送機構ぎふMVLリサーチセンター研究員. 没入型ディスプレイおよび高臨場感通信に関する研究に従事. 修士(工学).

栗田裕二

(正会員)



1990年早稲田大学理工学部電気工学科卒業. 1991年NECソフトウェア中部入社. 1999年より通信・放送機構ぎふMVLリサーチセンター研究フェロー. バーチャルリアリティの研究に従事.

服部陽一



1989年中部大学工学部卒業. 同年日本電信電話株式会社入社. 1999年NTTコミュニケーションズ株式会社に移籍. 2000年より通信・放送機構ぎふMVLリサーチセンター研究フェロー. 仮想空間を利用した人体の立体画像化の研究に従事.

廣瀬通孝

(正会員)



1977年東京大学工学部産業機械工学科卒業. 1982年東京大学大学院工学系研究科博士課程修了. 同年東京大学工学部講師. 1983年同助教授. 1999年東京大学先端科学技術研究センター教授. システム工学, ヒューマンインタフェースの研究に従事. 工学博士.