

デスクトップVR会議におけるアバタ表現方式の評価

松本 敏宏, 松浦 宣彦, 菅原 昌平, 正木 茂樹
NTTサイバースペース研究所 メディア通信PT

ネットワーク上に構築された共有仮想空間であるインタースペースの中では、アバタと呼ばれる自分の分身を介して、同じ空間に参加している他の分身とコミュニケーションをはかることができる。しかし、アバタの顔の部分にのみ自分の顔画像をまとめた。現在のアバタの表現方式ではその人の印象を伝えることができない。そこで、我々は仮想空間でのより効果的なコミュニケーションの伝達方法を模索するために、今回はデスクトップ上の仮想空間を用いた簡易会議という用途に限定して、実映像とコンピュータグラフィック (CG) をどのようにアバタに組み合わせればコミュニケーションに効果的であるかという評価実験を行なった。その結果、デスクトップ上の簡易会議という用途に対しては、上半身レベルの実映像とCGを組み合わせさせたアバタの表現方式が効果的であるとの認識を得た。

Evaluation of an expression for the avatar on Desktop Virtual Reality

Toshihiro MATSUMOTO, Norihiko MATSUURA, Shohei SUGAWARA, Shigeki MASAKI
NTT Cyber Space Laboratories Media Communications Project

We can communicate with each other through the avatar on networked virtual environment. We can't impress others well, for present method of avatar expression through only real facial video. So, we want to find effective way of communication, therefore, for the purpose of desktop simple conference, we have examined the way which effectively combine real video with computer graphics for the avatar. The consequence shows that the combination of real video from the waist up level and computer graphics is effective for the avatar expression.

1 はじめに

インターネット等の通信手段で画像や音声、テキストチャット等によるメディアを駆使してリアルタイムに地理的に離れた相手とコミュニケーションをとることができるシステムが次世代のコミュニケーションシステムとして注目されている。例えば、遠隔会議システムやチャットによるコミュニティ等が挙げられる。しかし、遠隔会議システムとしては複数の人が参加できる場合、高価なシステムとなってしまう。また、チャットシステムとしてはエンターテインメント目的に限られたものが殆どで、そのためネットワークの向こう側にいる実際の相手をさほど意識しなくても良いので気軽に使用される。しかしその反面、相手が見えないので自分を偽ることも容易に出来てしまうという問題点もある。

この様に、ネットワークを利用して距離・時間を越えた新しいコミュニケーションの形態が今後ますます広まろうとしているが、一般家庭にまで広く提供可能なレベルに至っているシステムは未だ少ない。

著者らは、ネットワーク上に構築された仮想空間上に自らがアバタと呼ばれる分身になって参加し、同じく参加している人同士で音声、顔の動画像、アバタによる簡易なジェスチャやテキストチャットによる複合されたメディアを通してコミュニケーションが図ることのできるシステム、インタースペースの研究開発を行っている。特長は、インターネットに接続できる環境とマイク、カメラ等と専用のブラウザさえ準備できれば、誰でも容易に3次元仮想空間で、実際の相手を意識した

コミュニケーションに参加できることにある。これまで、主にシステムの安定性や一般的なPCでの動作、通常のダイヤルアップインターネット環境での利用、データベースシステムとの結合等コミュニケーションに付随する機能拡張重視で開発を進めて来た。しかし、一方で用途にあった空間デザインやユーザインタフェースの機能向上についてはあまり着目して来なかった。そこで、著者らはインタースペースを1つのビジネスモデルに適用するために、容易にインターネットの向こう側にいる複数の相手と同時にコミュニケーションがとれるという特長を活かして、デスクトップ上の簡易会議システムとしてシステムを利用する際の機能評価を行なうことにした。特に、

- ・コミュニケーションのし易さ
- ・ネットワーク帯域を考慮したデータ量
- ・デスクトップという制限された空間

という課題に対して、仮想空間の中で自分を演じるアバタ表現に着目し、最適な表現方法とは何かを検証していく。

このような3次元共有仮想空間を用いたシステムでのコミュニケーションの効果の度合いをはかる試みは、中西[1]らの従来のデスクトップ会議システムと3次元仮想空間を用いた場合のデスクトップ会議システムでの会話の活性度の違いをはかった考察がある。ここでは、3次元での会話環境の方が会話が活発になると結論付けられている。また、渡辺[2]らのアバタのノンバーバル情報や生体情報を収集・解析し、体系的にコミュニケーション特性を解明しようという試みもなされている。その中で、アバタにその操作者の動きに応じて頭部の運動を与えるだけで、自己と他者の認識が容易に行なえるようになると述べられている。そこで、著者らはまず第1段階として相手の認知度という観点に立って、アバタ表現の手法について検討する。特に、ジェスチャパターンへの知見を得るために、またアバタモデルに如何に実写を用いれば効果的か

ということについての検証をすることにした。本論文では、1対1で行なった今回の検証結果から今後の本格的な多人数参加型のシステム構築へ向けての課題抽出を行なう。

2 高精細インタースペース：HiQIS (High Quality InterSpace)

これまで、特にインターネットに対応したインタースペースにおいては、ネットワークの帯域を考慮して機能を制限している。特に、現在一般的なPCの性能を考慮し、3次元CGのレンダリング負荷を極力下げるために、仮想世界およびそこに出現するアバタのCGモデルを極力ポリゴン数、テクスチャ量を少なくし、PCに与える負荷を軽減している。また、顔画像通信においてもそのフレームレートを下げたり、音声に関しても、サーバ側でミキシングをする等クライアントにかかる負荷をできるだけ軽減している。

そのため、現在のインタースペース一般公開実験[3]を通してユーザからは、アバタの表現に関する不満が一番多い。

しかしながら、PCに関しては、CPU処理能力、さらに3次元CGに対するアクセラレーション能力の進歩には目を見張るものがあり、一般的なPCでの処理可能な3次元CGモデルの品質は飛躍的に向上している。

そこで、昨年度開発した「高精細インタースペース：HiQIS(High Quality InterSpace)」[4]では、アバタ表現の向上に重点を置いて、以下の改良項目を挙げた。

- (1) アバタをより人間らしくするための機能

- ・アバタのCGモデル高精細化
- ・アバタの動作モデルのポーズ増加
- ・首のみを動かす動作の提供

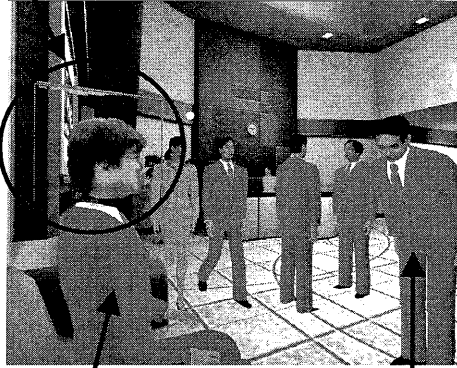
- (2) 顔画像の高精細化

- ・クロマキー合成による実写顔映像のアバタへのシームレスな統合

以下に、HiQISのユーザC(ユーザAとB

が会話中) が使用している PC の画面例を挙げる。

ユーザの右顔をリアルタイムに貼りつけ
(ユーザ A がユーザ B 方向に向いているため横顔が表示されている)



ユーザ A アバター
(会話モード)
ユーザ B と会話している

ユーザ B アバター
(歩行モード)
ユーザ A と会話している

図1 HiQIS 画面例

これらの機能により、

- ・アバターが歩行する際に人間らしくなった
- ・首振り機能によりアバターがどちらの方向を向いていて、空間内の何を、見ているかが特定できるようになった
- ・操作者毎のモデルを準備することにより、離れた所でもユーザを特定できる
- ・会話モードではアバターの CG モデルの顔部分に実写顔画像を使用しているため、相手の雰囲気がかめること等により、コミュニケーションする相手の認知という点においてこれまでのインタースペースより大幅に向上した。

しかし、ジェスチャのパターンも限定されており操作者の意思に応じて自由にジェスチャができない、また実写顔画像とアバターの CG モデルとの接合が困難であり不自然に見えてしまう場合があった。さらに、会話モードではアバターの CG モデルは座ったままで動作しないので相手の雰囲気伝達という点にお

いても多少違和感があるという問題点も分かった。その他、このシステム作成にあたっては、クロマキー合成のためのブルーバック、操作者撮影のためのマルチカメラの配置等、一般的なオフィス環境で実現するのは困難であるという問題点もある。そこで今後は、デスクトップ上という限られた空間の中で最適なコミュニケーションを行なうために必要なアバター表現については、そのジェスチャ機能について検討を進めて行く。

3 アバター表現方法の検討

デスクトップ上での仮想空間での最適なアバター表現について検討をするにあたって、HiQIS の問題点としてアバター動作の違和感が大きいことが分かった。

そこで、ジェスチャ機能を実現するアプローチとしては、以下の2つが考えられる。

(1) ジェスチャ機能の拡充

- ・ジェスチャ認識を利用してそれをアバター動作の指示に反映させる方法
- ・アバターの骨格構造を定義して各関節の動きデータのみを相手に送信して動きを持たせる方法

(2) 実写映像の拡充

- ・上半身等顔画像以外の体の実写映像の利用

(1) のジェスチャ認識に関しては現在活発に研究がなされている。また、アバターの骨格構造の定義とその動きデータの送受信による、アバターのリアルなジェスチャの実現方法は榎尻ら [5] により開発されている。(2) の手法については、会議や秘話などの場面においては実写を使用する方法といったような、場面に応じて実写映像と CG モデルとを組み合わせるといったものである。今回、我々は、インタースペースのデスクトップ会議システムへの適用という観点に立って議論を進めているので、(2) のアプローチが有効であると考えた。次節以降では、(2) のアプローチの有効性を確認するための評価システムの検討を

行なう。

4 システムの構築

4.1 システム構成

実際に制作した評価システムの構成について説明する。システムの構築にあたって本論文では、まずマルチユーザ対応ではなく、1対1でのコミュニケーションの実現から始めることにした。これは、

- ・ 1対1でのシステム構成であれば、インタースペースブラウザとクロマキー合成した映像とを別々のレイヤで重ね合わせることによる合成が可能であり、構築が容易である
- ・ 1対1の場合で、ある程度の知見が得られれば、マルチユーザ化の重要な指針となるなどの理由による。マルチユーザ化するための課題については後述する。以下に、実際の構成図を載せる。

実際のシステムは、図2のパターンがもう1セット存在する。

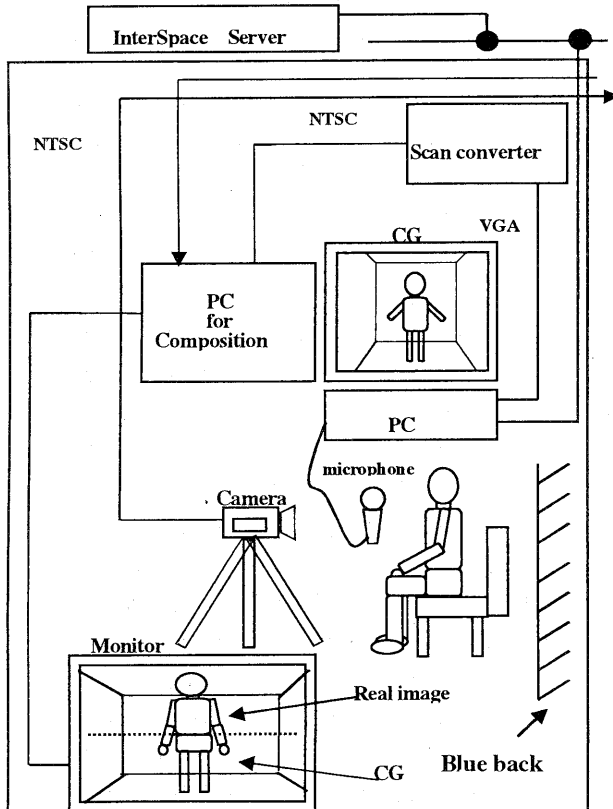


図2 システム構成図

先程触れた様に、インタースペースブラウザによる会話相手のアバタが表示された CG の画面と、相手のカメラによる画像は一度自分の PC の映像合成用ボードに NTSC の信号として入力される。相手のカメラの画像を直接ボードに入力したり、CG の画面をスキャンコンバータを通して VGA の信号から NTSC の信号に変換するのは、合成処理用のボードが NTSC の信号のみしか受け付けなためである。こうして、インタースペースブラウザの CG の画面と相手の実写映像とが別々のレイヤで重ね合わせるにより合成され、相手アバタに対して実写の映像と CG モデルとが合成されてモニタに表示される。この時、ユーザはインタースペース用に使用されている PC に接続されているマイクと最終的に合成して出力されたモニタを見ながら、相手と会話することになる。

4.2 システムの評価実験

前述のように構成したシステムで以下の項目で、被験者5組に対してその主観的な評価をしていただいた。

それぞれ、アバタが立位と座位のパターンで

- ・ 全身 CG
- ・ 顔の部分のみ実写でその他 CG
- ・ 上半身のみ実写その他 CG

の場合についてそれぞれある自由な話題について会話をしていただいて、会話のし易さ、相手の認知度、違和感等の観点で5段階評価をしていただいた。また、アバタに必要なと思われるジェスチャに関しても自由に記入していただいた。以下、その画面例の1部を載せておく。



図3 実写映像とCGモデルとの合成例

5 考察と今後の展開

今回の使用者による主観的な評価に基づき、改善すべき点やマルチユーザ化に向けた今後の進め方を検討する。

5.1 インタフェースの改善

一番多かった意見として、今回のシステム構成の都合上アバタに実写映像を重ねる際に非常に労力をさかなければならないということが挙げた。相手のモニタに自分の姿が映ってはいないものの、自分では姿は見えないので、会話相手に、先導してもらいながらの重ね合わせの作業となってしまった。また、合成されたとしても、実写とCGとの接合が不自然に感じられるという意見も多かった。

以上の様な問題点はポリゴンに実写映像を貼りつけるシステムにおいて、接合部分の位置調節を自動に行なう機能があればある程度緩和されるものと考えられる。

5.2 アバタジェスチャについて

アバタのジェスチャのパターンについては今回の実験においては、十分なジェスチャパターンを準備できなかったのであまり評価することは出来なかった。

現在のジェスチャの仕組みは、マウスを用いて仮想の空間内を移動する(Walk Throughするという)際にそれを見ている相手には、あたかも歩いているように見せるというものであったり、相手アバタをマウスでクリックして、お辞儀ボタンを空間内に表示させてから、さらにそのボタンを押すことによりアバタが実際の動作に入るといった仕様になっている。しかし、このようにジェスチャをするためには空間内であるイベントを操作者自らが発生させなければ出来ないような仕組みになっている。また、システム構成も複雑になっている。また、システム構成も複雑になっている。また、システム構成も複雑になっている。これがユーザにシステムへの不満を与える大きな要因になっていることが改めて明確になった。

5.3 実写映像との合成について

(1) 全身CGの場合

まず、全身CGの場合はやはり、そのアバタのデザインにより大きく左右される。今回の実験システムでは、当初ある一定のジェスチャパターンを持ったアバタを実現するために様々な種類のCGモデルを準備しなければならなかった。そのため、アバタにテクスチャを貼る作業までには至っていなかった。しかし、そのアバタを使用して実際に画面に表示させた所、これまでよりも相手を認識する点においてかなり違和感がでてしまった。そこで、HiQIS(図1を参照)で使用していたアバタに切り替えた所、相手の認識度は見違えるように良くなった。この要因はHiQISで用いられているアバタのCGモデルのテクスチャによることが分かった。HiQISではCGの顔のテクスチャは実際の人物の顔の写真を基に作成されている。そのため、かなりリアルである反面、その作成にはコストも時間もかなり要する。また、CGモデルの体の部分は全てのアバタで同じテクスチャ、同じ動きをするので、動きの部分がその操作者毎に異なればその認識度はさらに一層高まるものと考えられる。

(2) 顔の部分のみ実写でその他CGの場合

これに関しては、前述の通り、CGモデルと実写顔映像との接合が上手くいかない場合が多く、また会話モードになっている場合でもアバタモデルは動かないので会話中にも多少の違和感を感じた被験者が多かった。これには、自動に実写映像の位置・大きさを調整する機能が備わっていれば多少違和感は緩和されるものと考えられる。

(3) 上半身のみ実写でその他CGの場合

この場合は、上記2つのパターンより一番コミュニケーションし易いという意見が多かった。この場合、立位の場合では、やはり上半身の実写映像と下半身のCGモデ

ルとの合成で違和感があったが、立ち話や会議の様に椅子に座った場合、下半身 CG モデルは殆ど見えず、上半身の実写のみが見えた。会話し易いと回答していただいた方が多かったのはそのためだと考えられる。

しかし、今回の実験では1対1の場合であったが、これをマルチユーザ化するためには、大きな課題がある。現在の HiQIS 版では、実写顔映像は、正面・左右の3方向からのカメラ映像を仮想の視線の方向に合わせて切り替えている。そのため第3者から見ても、仮想空間の中で顔の CG モデルが回転しているのが伺えるし、会話モードに入っているアバタ同士でも顔の映像が違って見ることができる。これを、今回のシステムの延長として実現しようと試みるには、ある2人がコミュニケーションモードに入っていて、そこに第3者が加わった場合にはその表示方法は非常に困難である。現時点では、正面から映像をならべて表示するか、四方八方にカメラを設置しておいて、仮想の視線に合わせて撮影するカメラを自動で切り替えるという様な方法が考えられ得るが、前者の方はデスクトップ上という場所の制限があり、後者の方は実用的ではない。

5.4 ネットワーク帯域への課題

5.1 から 5.3 章で述べたことを踏まえて、今後のマルチユーザ化に向けてのネットワーク帯域への課題を挙げておく。現在の HiQIS 版では実写映像はマルチキャストによりネットワーク上に常に流れているが、100Mbps の LAN を使用しているので、数十名程度まで使用可能である。しかし、今後 10Mbps 程度で動作可能になることを考えると、コミュニケーションモードになって始めて映像を送信する等の LoD を考えていかなければならない。

6 まとめ

本論文では従来のインタースペースで用いられていたアバタ表現では充分なし得なかつ

たコミュニケーション方法を見直すための知見を得ることを目標に、1対1の場合での実写映像との最適な組み合わせの実験を通して調べた。特に、仮想空間の中で椅子に座って話す場合やある程度近づいて話しをする場合は上半身の実写映像が表示されると相手の雰囲気良く伝わり、コミュニケーションし易いという結果が得られた。しかし、その反面それをマルチユーザ化しようとする、現時点で考えられるシステム構成では一般化は困難である。

そのため、デスクトップという制限を十分に生かすためにできるだけ CG で表現し、秘話や少数の会議では実写を使用するといったような、コミュニケーションのモードによって実写映像と CG とを切り替えるというのが現時点での最適な方法だと考えられる。また、アバタのジェスチャ方法も現方式とは違ってある程度のパターンを操作者個人に応じて変更できる仕組みの実現が必要である。今後は、その2点に絞ってさらに評価を進めて、システムの実現を目指して行く。

7 参考文献

- [1]中西英之, 西村俊和, 石田亨, “デスクトップ会議における3次元仮想空間の効果”, 情報処理学会論文誌, Vol.39, No.10, pp.2770-2777(1998)
- [2]渡辺富夫, 大久保雅史, “身体的コミュニケーション解析のためのバーチャルコミュニケーションシステム”, 情報処理学会論文誌, Vol.40, No.2, pp.670-676(1999)
- [3] <http://CyberSociety.elcs.intsp.or.jp>
- [4]菅原昌平, 松浦彦彦, 箕浦大祐, 松本敏宏, 正木茂樹, “高精細型インタースペースによる会話環境の検討”, 情報処理学会 GW 研究会, 99-GW-31, pp.67-72(1999)
- [5]樋尻俊紀, 西谷和博, 望月義幸, 中俊弥, 浅原重夫, “VRML ヒューマノイドアニメーション用動作ストリーム符号化の一手法”, 情報処理学会グラフィクスと CAD 研究会, 91-8, Vol.98, No.76, pp.41-46