

実世界情報システムプロジェクト～バーチャルリアリティ研究グループ～

実空間アバタと人の新しいインタラクション に関する研究成果

廣瀬通孝 広田光一
先端科学技術研究センター 生命知能システム分野

概要

実世界VRをコミュニケーションに利用する際に不可欠な要素である実空間ビデオアバタ技術を確立し、日常的環境における臨場感の高いコミュニケーションの実現をめざす。本報告では今年度に検討した実世界ビデオアバタの2つの提示手法について述べる。

1 はじめに

近年、ビデオアバタを用いた臨場感の高いネットワークコミュニケーションに関する研究が多数行われている[1]。しかしながら、それらは没入型ディスプレイのような特殊な装置を用いるものであり、VR空間に限定されたものであると言える。このような反省からオフィスなどの実空間にVR環境を展開する実世界VRの考え方が提案され、これをコミュニケーションに利用する試みが行われるようになってきた[2]。また、人と環境とのインタラクションについても、VR空間の情報を実世界にあるものと同様の操作で扱うことを可能にする実世界指向のインタフェースが提案されている[3]。

一方、VR空間をコミュニケーションに利用する際には遠隔の人物を現実感や存在感をともなって提示することが必要とされ、その実現方法としてビデオアバタと呼ばれる実写ベースの人物像生成の研究がおこなわれている[4]。ビデオアバタを提示することでその空間内におけるユーザー同士の位置関係や姿勢などの空間的な情報が伝達できるようになるが、これまではVR空間での利用が想定されており、実空間でのビデオアバタの提示については検討されていなかった。

このような背景をふまえて、本研究では、実世界VR環境におけるビデオアバタの提示技術に関して2つの手法を検討した。ひとつはプロジェクタにより人物像を環境に投影する実世界ビデオアバタ投影による手法であり、もうひとつは実

空間中で視点依存の映像提示をおこなうことのできる実世界ビデオアバタディスプレイを利用する手法である。

2 実世界ビデオアバタ投影

2.1 システムの構成

本研究においては、アバタを壁面に投影するための可動型プロジェクタとして東芝ライテック社のActiveVision(図1左)を用いた。この装置は、プロジェクタ本体を上下左右に自由に回転することで壁面の広範囲に映像を投影可能であり、DMX信号によりズームや焦点調節などの制御も行うことができる。実験環境にはパーティションにより構成された壁面に加えて、半透明スクリーン(図1右)を設置することで、壁面でない実空間内部にアバタを表示する効果も検討する。



図1 ActiveVision と半透明スクリーン

2.2 画像の投影手法

実空間においてコミュニケーションを行う際には、立体視を行うための道具を使わずに壁面に投影されたアバタ映像を裸眼で観察することになるが、このように両眼視差を制御できない条件では、3次元形状の表現が難しいことが明らかになった。具体的には、プロジェクタの光源(以下、投影中心)と壁面との位置関係、観察者の視点位置

や視線方向などをもとに3次元形状を空間幾何学の理論に従って正しく投影したとしても、人は両眼立体視により壁面の三次元形状そのものを認識してしまい、壁面に投影された映像は壁面上の模様として認識されてしまうということである。この知見を踏まえて、本手法では壁面上の2次元映像を正しく再現するアプローチを採ることとする。すなわち、ポスターのように視点が正面にあることを前提とした映像を壁面の任意の位置に投影する。テレビなどを斜めからみても違和感がないのと同様に、この方法は主観的には違和感が少ない。このような投影の例を図2に示す。壁面上に投影された画像の理論値からの誤差は横：約3%、縦：約2%であった。

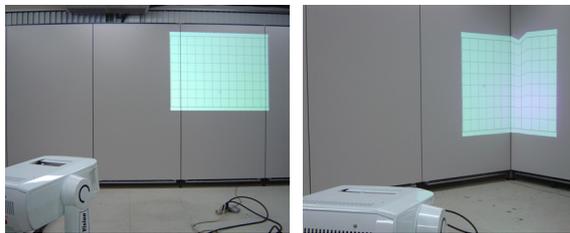


図2 投影例(左図:一壁面 右図:二壁面)

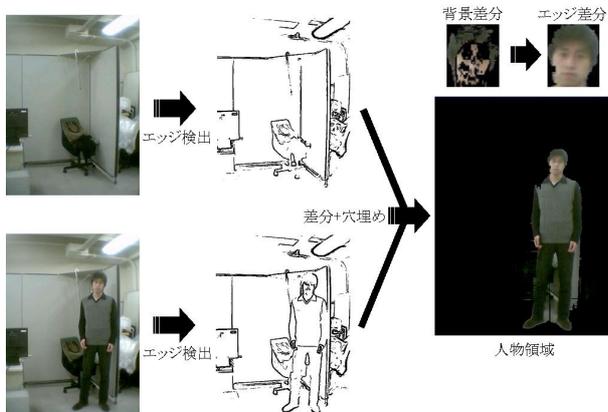


図3 アバタ画像の作成

2.3 アバタ画像の生成と提示

実空間でアバタ画像を作成するにあたり、クロマキー環境を用意することは難しい。また、背景差分による方法では、コミュニケーションにおける重要なノンバーバル情報である顔の表情がノイズにより十分に再現されないことが明らかになった。そこで本研究では、画像のエッジに関する差分を利用することで人物の輪郭線を抽出し、輪郭線の内側を人物領域として抽出することでアバタ画像を作成することを考えた(図3)。背景差分と比べて余計な領域が抜き出されているも

の、顔の表情は背景差分の場合よりもよく(約+10%)再現されている(図3右上)。

アバタの投影では、図4のようにプロジェクタを2台用いてアバタ画像を上半分と下半分に分けて投影することにする。これにより、1台で投影する場合の16/9倍の解像度を得る。なお、システムにおいてプロジェクタの位置は固定されておりプロジェクタの回転中心は固定であるが、回転中心と投影中心にはずれが存在し、投影の際に考慮する必要がある。このずれは、投影中心が回転中心に一致しているとして画像を変形した際に、壁面に投影される画像の理論値との誤差から求めることができる。なお、アバタ画像を投影する際、アバタが宙に浮いている印象を避けるため、足が床面に接するように投影した。また、アバタが壁面上を移動する際、上下の画像の移動速度を等しくする必要があった。上述の方法で作成したアバタ画像を2台のプロジェクタにより投影した例を図5に示す。

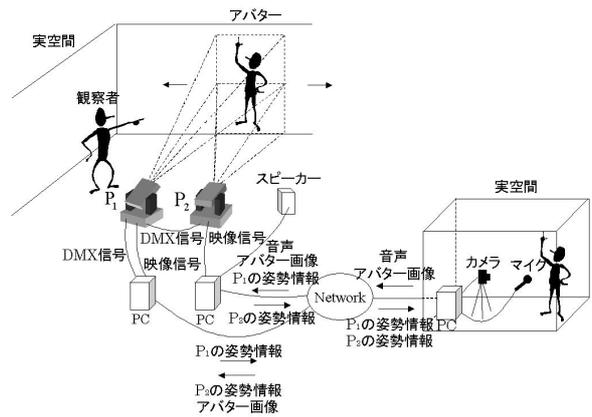


図4 システムの全体構成



図5 作成されたアバタ画像の投影

2.4 環境の評価

壁面にアバタを投影する際、アバタの全体がプロジェクタによって投影されること、アバタの解

像度が画像の変形により落ちないことの両方を満たす位置にプロジェクタを配置する必要がある。この点を考慮して、実験では画像解像度を横 240pixel×縦 320pixel、プロジェクタの回転中心と全ての壁面との距離がおおよそ 2.5m となるような配置とした。

一方、アバタ画像を上下別々に投影するため、上下の画像のずれを計測により確認した。これによると、横方向のずれ $\delta x < 1.0[\text{cm}]$ 、縦方向のずれ $\delta y < 1.0[\text{cm}]$ が達成され、主観的には実用上意識されない程度であった。

3 実空間ビデオアバタディスプレイ

3.1 アバタ提示手法

実空間に提示されるビデオアバタに高い臨場感を与えるには、映像の写実性に加え本物の人間のようにあらゆる角度から視認できることが重要である。それによりユーザとビデオアバタが自由な位置関係で、空間を生かしたインタラクションを取ることが可能となる。加えて実時間性も重要な要素である。全周囲から見ることのできる（全周視認性と呼ぶ）ディスプレイとしてはマルチブックスホログラムや体積走査型ディスプレイ [5] があるが、前者はホログラムの更新が困難であり、後者は空間にボクセルを描く方法であるため隠蔽関係を表現することができない。

高い指向性を持つ（すなわち、視認角度の狭い）視覚ディスプレイを用いることで特定の方向にのみ映像を提示することができる。このような高い指向性を持つディスプレイを高速で回転させ、その回転角に応じて映像を切り替えることにより、異なる映像を異なる方向に提示するというのが提案手法の原理である。複数台のカメラにより撮影された人物の全周映像をそれぞれ撮影された方向に提示することにより、視点位置の変化に応じて見え方が変化する 2 次元映像が提示される (図 6)。提示される全周映像のフレームレートは回転ディスプレイの画面更新レートをディスプレイが 1 周回る間に提示する映像の数で割った値となる。なお 1 回転で各方向に 1 回だけ映像を提示することから映像のフレームレートとディスプレイの回転数の値は等しいことになる。

3.2 システムの試作

人物の全周撮影には全周カメラシステム [6] を使用している。図 7 に全周カメラシステムの構成図を示す。18 台のカメラを円筒型の部屋の壁面に

等間隔で配置した構成である。カメラはそれぞれ 1 台の PC (ノード PC) に IEEE1394 経由で接続されており、またノード PC はギガビットイーサネット経由でシステムを統括する PC (統合 PC) に接続される。各カメラからは最大 15 フレーム/秒で XGA の動画像が得られ、その動画像はフレーム単位で処理され提示システムへ転送される。撮影された画像はノード PC において背景差分による人物像の抽出、描画速度の高速化のためのサイズ縮小、保存や転送の効率化のための JPEG 圧縮の処理がおこなわれる。画像は統合 PC からの要求に従い転送され、統合 PC は受け取った画像をさらに提示システムへと転送する。統合 PC はこの要求、取得、転送のプロセスを各ノード PC に対して順におこなう。ノード PC は他のノード PC に画像が要求されている間に画像の処理をおこなう。

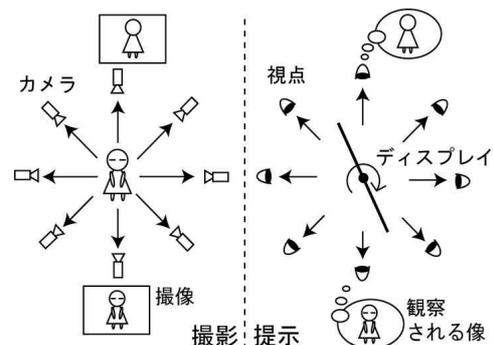


図 6 システム概念図

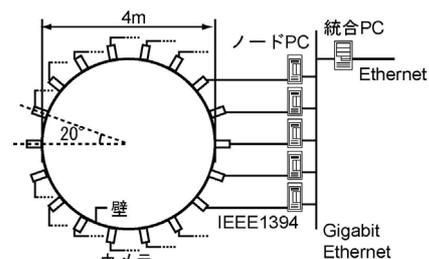


図 7 全周カメラシステムの構成

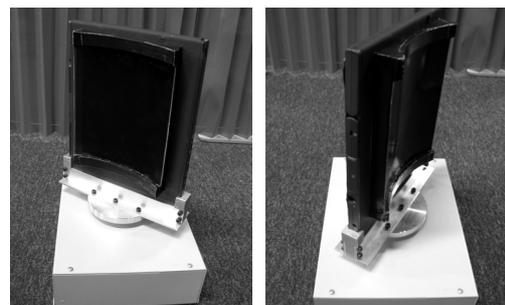


図 8 提示システム外観

提示システム(図 8)はタブレット PC(PC-TB7005T, NEC)とステッピングモータを用いた回転機構で構成され、約 1/10 スケールの人物像が提示される。ディスプレイは TFT カラー液晶で最大 60[Hz]で映像の更新をおこなう。ディスプレイには指向性を高めるためのフィルタが、カメラ位置に対応する視点からディスプレイ全面が一様に見えるように弧状(曲率半径:約 200mm)に曲げられた状態で取り付けられている。視点位置とフィルタ面との角度が大きくなるほど画面が暗くなりおよそ±24°で完全に見えなくなる。撮影システムからの画像は IEEE802.11a 準拠の無線 LAN 経由で受信する。画像は一定の時間間隔で描画され、回転速度はこの時間間隔に応じて設定される。

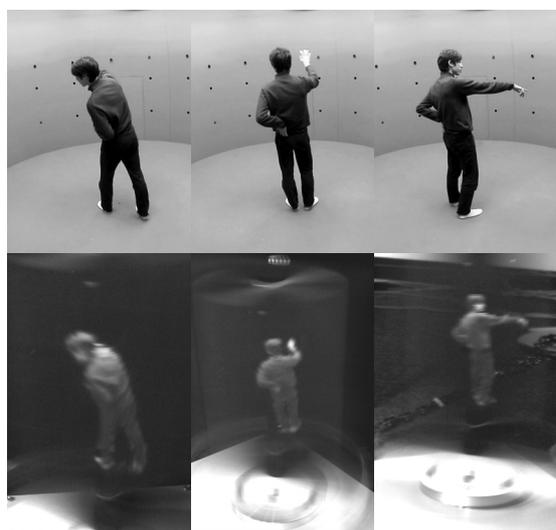


図 9 撮影(上)と提示(下)の様子

3.3 システムの評価

解像度 120x90 の画像を受信しながら描画をおこなう場合で 30[Hz]の安定した画像更新が可能であった。このデータを元に異なる映像を提示する方向を 6 方向、回転数を 5[rps]と設定した。すなわち、提示される映像のフレームレートも 5[Hz]となる。

提示されたビデオアバタの様子を図 9 に示す。上図は撮影された画像、下図はその画像が提示されている様子である。カメラによって撮影された映像が各方向に提示され、視点移動に従い映像が不連続的に変化する様子が観察された。一方、この実験から、フィルタの指向性が十分でないため異なる方向から撮影された 2 枚の人物像が重なって見えてしまう位置があること、また提示する環

境が明るいときと像が見えにくくなることなどの問題が、明らかになった。なお、撮影から提示までのトータルの遅延は平均 176[ms]であった。

4 おわりに

本報告では、実世界 VR 環境を利用したコミュニケーションを目的としてビデオアバタを実空間で提示する手法について述べた。実世界ビデオアバタ投影については、投影方向に制御の自由度をもつプロジェクタにより実空間において自由度の高い人物像提示が可能であることを明らかにした。これをふまえて来年度はアバタとユーザとのインタラクションの枠組みの検討やコンテンツ依存の提示手法の提案をおこなう予定である。一方、実世界ビデオアバタディスプレイについては、視点依存の映像提示のための新規な手法を提案し、その実装とビデオアバタ提示の可能性を検証した。このディスプレイの時間空間の分解能には改善の余地があり、今後検討していきたい課題である。

参考文献

- [1] 広田, 廣瀬, "マルチメディアバーチャルラボラトリ", 日本 VR 学会論文誌, Vol.8, No.1, pp.47-56, 2003.
- [2] G.Welch, H.Fuchs, R.Raskar, H.Towles, M.S.Brown, "Projected Imagery in Your Office of the Future", IEEE CG&A, Vol.20, No.4, pp.62-67, 2000.
- [3] P.Wellner, "Interacting with Paper on the DigitalDesk", Communications of the ACM, Vol.36, No.7, pp.87-96, 1993.
- [4] K.Tamagawa, T.Yamada, T.Ogi, M.Hirose, "Developing a 2.5-D Video Avatar", IEEE Signal Processing Magazine, Vol. 18, No. 3, pp. 35-42, 2001.
- [5] J.Napoli, D.M.Hall, R.K.Dorval, M.G.Giovinco, M.J.Richmond, W.S.Chun, "100 Million-Voxel Volumetric Display", AeroSense 2002 - for Cockpit Displays IX: Displays for Defense Applications, 2002.
- [6] 谷川, 広瀬, 山田, 小木, 広田, 廣瀬, "全周カメラシステムによる没入型ディスプレイのための 3 次元アバタの研究", 日本 VR 学会論文誌, Vol.8 No.4 2003, pp.389-397, 2003.