

JGN2 ネットワークを用いた マルチカメラによるビデオアバタ送信

Transmission of multi-camera based video avatar through the JGN2 network

酒井満隆¹⁾, 小木哲朗²⁾, 鈴木康広³⁾, 広田光一⁴⁾

Mitsutaka SAKAI, Tetsuro OGI, Yasuhiro SUZUKI and Koichi HIROTA

1) 筑波大学大学院 システム情報工学研究科

(〒305-0006 茨城県つくば市天王台 1-1-1, mitsu@gil.cs.tsukuba.ac.jp)

2) 筑波大学 学術情報メディアセンター

(〒305-8577 茨城県つくば市天王台 1-1-1, tetsu@cc.tsukuba.ac.jp)

3) 東京大学大学院理工学系研究科

(〒113-0033 東京都文京区本郷 7-3-1, thzuki@cyber.rcast.u-tokyo.ac.jp)

4) 東京大学大学院新領域創成科学研究科

(〒227-8563 千葉県柏市柏の葉 5-1-5, k-hirota@k.u-tokyo.ac.jp)

Abstract: In this study, the JGN2 network environment was constructed to transmit the video avatar data from the University of Tsukuba to the University of Tokyo. The user's video images were captured from many directions using the multi-camera system at the Tsukuba site, and they were displayed on the rotary multi-view display at the Tokyo site. This display system represents the video avatar that can be seen from many directions in the real world. By using this system, the communication experiment was conducted and the performance of the video avatar transmission was examined.

Key Words: JGN2, video avatar, multi-camera, rotary multi-view display.

1. はじめに

近年、遠隔地とのコミュニケーションにおいてビデオアバタの利用が注目されている[1]。この方法は、人物の表現にCGではなく実写映像を使うことで、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。ビデオアバタの表現方法としては、これまで2次元モデルや3次元モデル[2]を用いた種々の方法が提案されている。遠隔地の人物を3次元映像としてVR環境に表現する方法は、より自然なコミュニケーションを実現することができるが、実時間での通信や表現を考えるとまだ処理時間に問題がある。

一方、VR空間の共有システムでは没入型ディスプレイを用いる研究が行われている。ビデオアバタを没入型ディスプレイで使用することで実物大の人物投影を可能にすることができる。また、没入型ディスプレイのような特殊なVR環境ではなく、実空間の環境を利用してビデオアバタを提示する研究も行われている[3]。

本研究では撮影環境として、マルチカメラを使用し人物を複数の角度から撮影したビデオアバタデータを生成するシステムを構築した。また、ビデオアバタの描画に液晶パネルを回転させることで実空間での全周囲視認性を実

現するシステムを構築した。更にビデオアバタ生成と描画を別々の地点で行うことを可能にするため、JGN2ネットワークを利用したビデオアバタの通信環境を構築し、実時間性のパフォーマンスに関する検討を行った。

2. ビデオアバタ

2.1 人物の切り抜き

ビデオアバタを遠隔地の世界に合成表示するためには、人物のみの映像を必要とする。リアルタイムで人物以外の背景領域を除く代表的な手法として、ブルーバックを用いる方法がある。この方法はクロマキーを用いることで、比較的きれいな処理を行うことが可能であるが、背景条件を特殊な環境に整えなければならないという問題がある。

そこで、本研究では背景差分法を用いて全身の実写人物像をリアルタイムで獲得する方法を用いた。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撮影映像の差分から目的とする人物像を抽出する方法である。図1は、1台のカメラを用いて背景画像と観測映像から人物だけを切り抜いた画像を示したものである。



図1 背景差分の実行例

2.2 マルチカメラ環境

人物を撮影する環境としては、ビデオアバタに高い臨場感を持たせるためにマルチカメラシステムを構築した。マルチカメラを用いることで人物を複数の角度から撮影することが可能となる。図2にマルチカメラシステムの構成を示す。これは、人物の周囲を8台のカメラで取り囲む全周囲システムで、それぞれのカメラは直径4mの円周上に45°間隔で均等に配置している。各カメラには、IEEE1394カメラであるSONY DFW-X710を用い、8台のノードPC (CPU: Pentium4 3.0GHz、OS: 4台がRed Hat Enterprise Linux、残り4台がFedora Core4)とIEEE1394で1対1に接続している。映像は解像度1024×768[pixel]、最大15[fps]で画像1フレーム単位に獲得することができる。各ノードPC8台では画像1フレームを獲得すると同時に、前述の背景差分処理を行う。

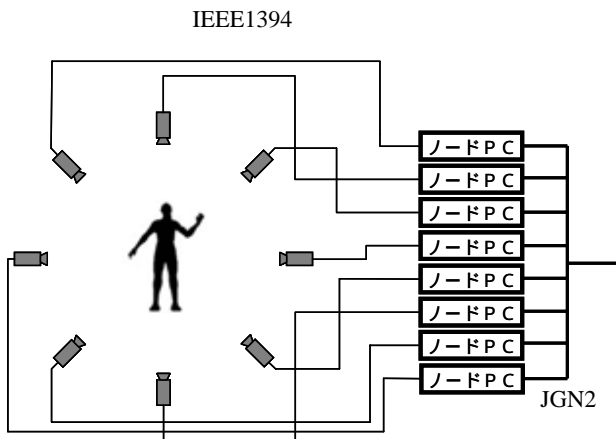


図2 マルチカメラシステム構成図

本研究ではビデオアバタの通信を目的としている。そのため、背景差分処理が終わると画像を1フレーム単位で遠隔地にあるPCに送信する。

2.3 ビデオアバタ通信

今回の実験では筑波大学と東京大学の間で通信を行った。全周囲視認性のために8台のカメラで撮影された映像を実時間で送信するためには、広帯域のネットワーク環境を必要とする。本研究ではつくばWANを経由してJGN2 (Japan Gigabit Network 2)を使用して、筑波大学 - 東京大学間を接続するネットワーク環境を構築した。JGN2は、情報通信研究機構が運用する研究開発テストベッドネットワークで、最大20Gbpsのバックボーンを有している。また、つくばWANは筑波研究学園都市の研究機関を10Gbpsのアクセスリングで結ぶ高速ネットワーク環境である。

通信の画像データ形式にはJPEG等に圧縮する方法も考えられるが、今回の実験では全体のパフォーマンスを考え非圧縮型のRGB形式の画像データを用いた。通信には1回の送信に関して最大1024×768の解像度の画像データを15[fps]で送らなければならない。さらにマルチカメラを使用しているため、同時に8台分のデータを扱うことになる。表1に映像の解像度とデータサイズについての関係を3通り示す。

解像度	1024×768	640×480	320×240
データサイズ	2264.9	884.7	221.2

単位: M bit

15[fps]でカメラ8台分のデータを送信するには、解像度1024×768で約2[Gbps]の帯域を必要とすることがわかる。また、解像度を640×480、320×240とすればギガビットネットワークの環境上でマルチカメラシステムを用いても、ビデオアバタデータの実時間送信は可能な計算結果であることがわかる。

3. 回転型多視点ディスプレイ

3.1 回転型映像提示

本研究ではマルチカメラ映像から3次元のビデオアバタを構成するのではなく、2次元データのまま3次元的な表示を行う方法を用いる。ビデオアバタの提示装置としては、回転型多視点ディスプレイを使用した[3]。この装置は、特定の方向にのみ映像を提示することができる高い指向性を持つディスプレイ面を高速で回転させ、回転角に応じて映像を切り替えることで、目的の映像を目的の方向のみに並列に提示することが可能である。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、2次元映像を用いながら視点位置の変化に応じた追従映像を提示することができる。

3.2 映像提示システム

映像提示システムとしては、2台のタブレットPC(NEC: PC-TB700ST)を用い、表裏2枚張り合わせて回転させる

映像提示システムを実装した。システムの構成を図 3 に、システムの外観を図 4 に示す。本システムは、タブレット PC と DC モータ (ORIENTAL MOTOR:BX-230A-A) を用いた回転機構で構成され、システム全体の高さ約 400[mm] で約 1/10 スケールの人物像が提示可能である。

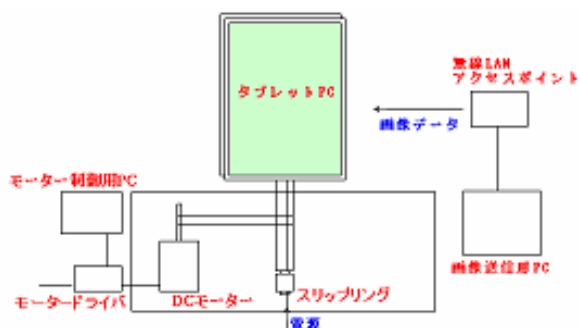


図 3 回転型多視点ディスプレイの構成



図 4 回転型多視点ディスプレイの概観

ディスプレイ面は TFT Color 液晶で構成され、最大 60[Hz]で映像の更新を行うことができる。表示面には映像の指向性を高めるためのプライバシーフィルタを貼ることでディスプレイの視野角を 48° に抑え、表示面以外は外部からの光の反射を避けるために黒く塗装した。表裏二枚構造となっており、表と裏のディスプレイ面に提示される人物映像は常に反対方向からの視点映像となるように同期を取っている。また HDD は回転や振動に弱いため、CF-IDE 変換アダプタを用い CF メモリカード(4GB)をタブレット PC の HDD の代わりとして使用し、回転や振動に強い構造とした。

DC モータはモーター制御 PC により回転速度を 0.1[rpm] 単位で制御可能であり最大回転数は 300[rpm]である。下部スリッピングは 250[V]、30[A]までの電源を 6 本供給可能であり、回転するタブレット PC へ 16[V]、4.5[A]の電源を供給する。これによって本提示システムはバッテリーに依存することなく連続稼働が可能となっている。

またタブレット PC は無線 LAN によって通信を行い、JGN2 に接続された画像配信用 PC からビデオアパ映像

を転送される。無線 LAN アクセスポイントは IEEE 802.11a を用いており、54Mbps によって通信を行っている。

3.3 予備実験

つくば - 東京間で実際にビデオアパの実時間通信を行う前に、回転型多視点ディスプレイ内で予備実験として画像のサイズによる映像配信速度を計測した。具体的には、画像配信用 PC に保存された jpg 画像ファイルを展開し、タブレット PC の回転に応じて対応する画像を配信した。画像サイズと通信速度の結果を表 2 に示す。この結果から、画像サイズに関わらず約 8.9[Mbps]での通信ができていることを確認できた。

表 2 画像サイズと通信速度

	120	140	160	180	200	240	320
幅	120	140	160	180	200	240	320
高さ	80	105	120	135	150	180	240
fps	34.2	25.3	21.8	15.0	12.3	7.95	5.30

4. つくば - 東京間通信実験

4.1 実験環境

つくば - 東京間の通信実験では、撮影環境として、2.2 節のマルチカメラシステムを用い、描画環境としては 3.2 節の回転型多視点ディスプレイを用いた。通信部分では、筑波大学のマルチカメラシステムの各ノード PC、東京大学に配置してある 1 台の通信用 PC(CPU: Xeon 3.0GHz, OS: Red Hat Enterprise Linux)を JGN2 で接続した。また、その通信用 PC と回転型多視点ディスプレイのタブレット PC は無線 LAN で接続した。

4.2 実験概要

各ノード PC では、カメラから解像度 1024×768(XGA)の画像データを 1 フレームごとに獲得する。背景差分後のビデオアパデータを筑波大学から東京大学へ送信し、回転型多視点ディスプレイに描画した際の、データ転送と描画におけるパフォーマンスを評価した。送信時の画像 1 フレームの解像度は XGA そのままの場合、XGA から 640×480(VGA)、320×240(QVGA)へそれぞれ変換して送信した場合の 3 通りの方法について比較を行った。

ビデオアパの提示には 3.3 節の予備実験より、8 方向のカメラ画像を受信後、解像度を 160×120 に変換しタブレット PC に配信することにした。モータは 3[回転/sec]で回転させ、タブレット PC は回転に合わせて 24[fps]で画像を更新描画させたため、8 方向に対しそれぞれ 3[回転]×2[枚]=6[fps]でビデオアパ映像を表示することができる。

4.3 結果と考察

表 3 にビデオアパ送信時の各ノード PC8 台で得られたフレームレートの平均値、最大値、最小値を示す。XGA ではカメラでの画像獲得最高フレームレート値 15[fps]の 1/5 程度しか出ていないのに対して、VGA、QVGA ではほぼ 15[fps]、つまり JGN2 上で実時間でのビデオアパデータ送信が行われたことを確認できる。また、表 1 と表 3 の関係から今回の実験では個々の PC にかかるビデオアパ

生成の負荷よりも、ネットワークに通す画像解像度の大きさがビデオアバ通信全体のパフォーマンスに大きく影響していたと考えられる。

表3 アバタ転送フレームレート値

	平均			最大			最小		
	XGA	VGA	QVGA	XGA	VGA	QVGA	XGA	VGA	QVGA
ノード PC1	2.75	14.3	14.7	4.38	15.0	15.0	2.18	13.4	13.6
ノード PC2	3.18	14.7	14.6	3.67	15.0	15.0	2.55	13.4	13.6
ノード PC3	3.12	14.5	14.7	3.58	15.0	15.0	2.47	13.5	13.9
ノード PC4	3.14	14.9	14.7	3.67	15.0	15.0	2.46	13.8	13.8
ノード PC5	3.30	14.8	14.7	3.87	15.0	15.0	2.77	14.3	13.8
ノード PC6	2.87	15.0	14.8	5.43	15.0	15.0	2.30	14.6	14.3
ノード PC7	2.85	14.5	14.6	3.58	15.0	15.0	2.37	13.6	13.8
ノード PC8	3.09	14.7	14.8	4.75	15.0	15.0	2.51	14.0	13.9
全体	2.97	14.7	14.7	5.43	15.0	15.0	2.18	13.4	13.6

単位：fps

XGA では画像送信のフレームレート結果からもわかるように、1秒間に2、3回の映像更新しか行われていない。これに対しVGAとQVGAでは画像転送のフレームレート値が十分な値を得ることが出来ており、回転しているディスプレイのある1視点に注目した場合、毎回映像が更新されていることがわかる。

描画部分では、ビデオアバタの解像度がどの場合でもタブレット PC1 枚についての描画リフレッシュレートが24[fps]に固定されていたことから、タブレット PC1 枚について8方向に対しそれぞれの方向に3[fps]で描画できていたことになる。すなわち、ディスプレイ全体として見るとタブレット PCを2枚用いているため、1方向について6[fps]の表示を実現することができた。

表4 転送解像度と時間遅れ

解像度	XGA	VGA	QVGA
時間遅れ	0.84	0.50	0.31

単位：秒

また、表4はビデオアバタを撮影した時刻と描画した時刻における時間遅れを示したものである。計測方法としては、1/1000秒表示のタイマ映像にNTTの時報を同時録音し、ビデオアバタデータとともに送信し、撮影側と表示側で同時刻に表示されているタイマのカウンタを比較した。表では、各解像度における1秒間の動作時の平均時間遅れを示している。この結果では0.3~0.8秒の時間遅れであったが、この要因としては、JGN2上での通信の他、無線LANの配信、ビデオアバタデータの生成、表示データの作成等による処理時間等が考えられる。

図5、6は実験の様子を表したものである。図5は筑波大学において、マルチカメラで人物を撮影した様子を示した図である。図6は東京大学において、回転型ディスプレイにつくばから送信されてきたビデオアバタを描画した図である。



図5 ビデオアバタ撮影の様子

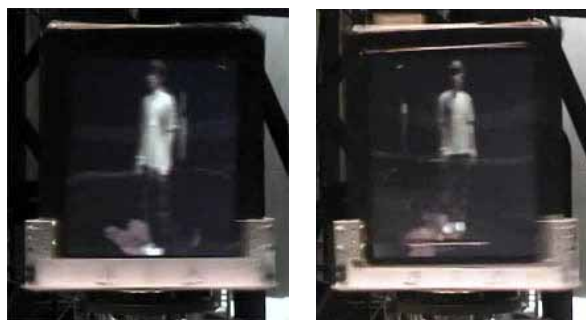


図6 回転型多視点ディスプレイへの描画

5. 結論

本研究ではマルチカメラと回転型多視点ディスプレイを用いることで全周囲視認性のビデオアバタ通信・描画を実現した。JGN2の特性を生かしてVGAの解像度まで実時間での通信を確認することができた。

実時間性を考慮した上でビデオアバタの人物切りぬきや画像解像度などの精度向上、ディスプレイの画像更新レート向上については今後の課題である。

謝辞

本研究は、総務省の戦略的情報通信研究開発推進制度「実空間における人物像の記録・伝送・再生に関する研究開発」として行われた。また情報通信研究機構のJGN2ネットワークおよびつくばJGN2リサーチセンターの研究施設を利用させていただいたことに感謝する。

参考文献

- [1] 小木, 山田, 栗田, 服部, 廣瀬: 仮想空間共有のためのビデオアバタ技術とその利用法, 日本バーチャルリアリティ学会論文誌, Vol.8, No.1, pp.37-46, 2003.
- [2] Stephan Wurmlin, Edouard Lamborary, Markus Gross: 3D video fragments: dynamic point samples for real-time free-viewpoint video, Computers & Graphics, 28,3-14, 2004
- [3] H.Maeda, T.Tanikawa, J.Yamashita, K.Hirota, M.Hirose: Real World Video Avatar: Transmission and Presentation of Human Figure, IEEE VR 2004, 237-238, 2004.3.