

Web ブラウザと全方位動画像を用いたテレプレゼンスシステム — 車載全方位カメラや全方位型マルチカメラシステムを用いた実装 —

石川 智也 山澤 一誠 佐藤 智和 池田 聖 中村 豊 藤川 和利 砂原 秀樹 横矢 直和

奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻

E-mail: {tomoya-i,yamazawa,tomoka-s,sei-i,yutaka-n,yokoya}@is.aist-nara.ac.jp, {fujikawa,suna}@itc.aist-nara.ac.jp

あらまし 近年、現実世界を仮想世界に取り込み、時間的・空間的に隔たった場所の事象を体験する技術であるテレプレゼンスの研究が行なわれている。従来、我々はテレプレゼンスにおいて臨場感を高めるために、利用者が提示されている映像の視線方向を自由に変えることのできる全方位動画像を用いたシステムを提案した。しかし、主な従来のシステムは没入型ディスプレイを用いたスタンドアロン型であり、一般的なネットワーク環境での実装には到っていなかった。そこで我々は、ネットワーク環境下で利用者が簡単に全方位動画像を用いたテレプレゼンスが可能なシステムを開発した。本報告では開発したシステムについて述べ、さらにリアルタイム型及び蓄積型の動画像を用いた実験について報告する。

キーワード テレプレゼンス, 全方位動画像, Web ブラウザ

Telepresence System Using Web Browsers and Omnidirectional Video Streams — Implementation Using Car-Mounted Omnidirectional Camera and Omnidirectional Multi-Camera System —

T.Ishikawa K.Yamazawa T.Sato S.Ikeda Y.Nakamura K.Hujikawa H.Sunahara and N.Yokoya

Nara Institute of Science and Technology Graduate School of Information Science

E-mail: {tomoya-i,yamazawa,tomoka-s,sei-i,yutaka-n,yokoya}@is.aist-nara.ac.jp, {fujikawa,suna}@itc.aist-nara.ac.jp

Abstract Recently, the telepresence which allows us to experience a remote site through a virtualized real world has been investigated. In previous work, we proposed the telepresence system using omnidirectional images which enable the user to see the virtualized dynamic scene in arbitrary direction. The system gives high presences to the user. However, the most conventional system is not realized by network oriented implementation but by stand-alone implementation that uses an immersive display. In this study, we have developed a new telepresence system which enable the user to see a virtualized real world easily in network environments. This paper describes the system in some detail and its experiments using real-time and stored video streams.

Keyword Telepresence, Omnidirectional video stream, Web browser

1. はじめに

近年、現実世界を仮想世界に取り込み、時間的・空間的に隔たった場所の事象を体験する技術であるテレプレゼンスの研究が盛んに行われている[1]。この技術はエンターテインメントや医療、教育など様々な分野に応用が可能である。

我々はテレプレゼンスにおいて臨場感を高めるために、利用者が提示されている映像の視線方向を自由に変えることのできる全方位動画像を用いたシステムを提案してきた[2,3]。これらのシステムは全方位画像センサにより遠隔地の周囲 360 度の全方位画像を取得・伝送し、その画像により利用者に視線方向の画像を提示する。この手法は利用者の視線変化からその方向の画像提示までの時間遅延が少なく、広範囲の遠隔

地を見回すことが出来るといった特徴がある。しかし、従来のシステムは没入型ディスプレイを用いたスタンドアロン型や専用プログラムなどを必要とするシステムであり、多地点で複数人が同時にシステムを利用するのは困難であった。

そこで我々は、ネットワーク環境下で複数の利用者が簡単に全方位動画像を用いたテレプレゼンスが可能なシステムを開発した。このシステムは Web ブラウザを利用し、画像などのコンテンツを楽しむのと同様に全方位動画像の鑑賞及び操作を行うことが出来る。同様なシステムには QuickTimeVR[4]や PanoVi[5]がある。しかし、QuickTimeVR は全方位静止画像のみを扱っており動画像には対応していなかった。また PanoVi は単眼カメラを組み合わせ、クライアント側でモザイクングを実時間で行うため、継ぎ目のない綺麗な画像を生

成するのは困難であり、画像の生成・提示に専用プログラムが必要で導入が困難という問題があった。

以下、2章では開発したシステムについて述べ、3章ではそのシステムに蓄積型及びリアルタイム型の動画像を用いた実験について述べる。最後に4章で、本稿のまとめと今後の展開を述べる。

2. 全方位動画像提示システム

本システムの概要を図1に示す。全方位動画像はネットワーク越しのサーバーやPC内に格納されており、Webブラウザにより起動されるビューアによってその全方位動画像を取得する。あるいは、カメラからの映像をリアルタイムに配信するサーバーからの全方位ビデオストリームを、Webブラウザにより起動されるビューアによって取得する。そして、利用者はWebブラウザ上でコンテンツを鑑賞・操作する。

2.1. Webブラウザを用いた全方位動画像ビューア

ネットワーク環境下で誰もが利用するアプリケーションにWebブラウザがある。特に、Windowsに標準で搭載されているWebブラウザであるInternet ExplorerはJAVAアプレットやActive-Xによるプログラムを実行可能で、利用者に対しWebページ上でインタラクティブなコンテンツを提供することを可能にしている。また、JAVAアプレットやActive-Xプログラムは自動インストール機能により利用者に負担をかけることなく提供可能である。

本システムにおいて全方位動画像を提示するソフトウェア（全方位動画像ビューア）はグラフィックスカードの機能を必要とするため、Active-Xプログラムとして実装を行う。全方位動画像ビューアはWebブラウザにより起動され、全方位動画像を平面透視投影画像に変換した後、Webページにその画像を提示する。この時、全方位動画像から平面透視投影画像への変換は尾上らの手法[2]によりハードウェアのテクスチャマッピング機能を用いて実時間で行う。

利用者は全方位動画像コンテンツが提示されるWebページにアクセスするだけで全方位動画像ビューアを自動的にインストールすることが出来る。また、全方位動画像ビューアに必要な全方位カメラの種類、カメラのパラメータ、動画像サーバーやファイルへのパス等はコンテンツ提供者がActive-Xプログラムを使用するWebページのHTMLファイルに書き込んでおけばよいので、利用者はカメラの種類等を意識することなくシステムが利用可能となる。

利用者は提示されている映像の上でマウスをドラッグすることにより、インタラクティブに視線方向の変更やズームイン・アウトなどを行うことが出来る。ソフトウェアの導入から操作まで利用者は指定された

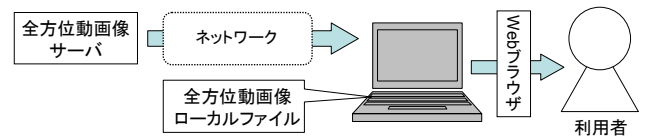


図1：全方位動画像提示システム



図2：全方位型マルチカメラシステム Ladybug

表1：蓄積型動画像観賞用PCの詳細

CPU	Pentium4 2GHz
メモリ	512MB
グラフィックスカード	ATI RADEON9700pro
OS	Windows XP

Webページを開くだけで行うことが可能で、簡単に全方位動画像を鑑賞可能である。また、姿勢センサとHMD (Head Mounted Display) を組み合わせることにより利用者に対し視線追従型の没入型画像提示を行うことも出来る。

2.2. 全方位動画像コンテンツ

全方位動画像は、あらかじめ動画像を符号化し、ネットワーク越しのサーバーやPC内に保存しておく蓄積型動画像とライブ中継のようにリアルタイムに画像を符号化し配信するリアルタイム型動画像に分けられる。本システムではその両方の動画像を扱うことが可能である。

蓄積型動画像では、ネットワークの負荷などの問題から伝送が難しいような高解像度の全方位動画像をPC内に格納しておき、鑑賞するといった目的に用いる。また、蓄積型動画像では利用者それぞれが、見たいシーンを見たい時刻から見るといったオンデマンドサービスが提供可能である。

リアルタイム型動画像は、TV放送のように同時に複数人が同じコンテンツを楽しむ場合に用いる。ネットワークの制限から高解像度の動画像を伝送するのは困難であるが、カメラにより取得された画像を瞬時に

伝送し鑑賞することが出来る。また、マルチキャストプロトコルを用いることによりネットワークに負荷をかけることなく多地点に伝送し、鑑賞することも可能である。

3. 実験

提案システムを実装し、蓄積型動画像とリアルタイム型動画像について実験を行った。蓄積型動画像を用いた実験では、全方位型マルチカメラシステムを用いて撮影された高解像度の全方位動画像を鑑賞することを目的とした。一方、リアルタイム型動画像を用いた実験では、全方位カメラなどを搭載した車両からの画像を無線・有線 LAN を通して伝送し、マルチキャスト配信によって効率的に複数人が同時に動画像を鑑賞することを目的とした。

A) 蓄積型動画像

全方位マルチカメラシステム Ladybug (図 2) [3]により得られた動画像を鑑賞用 PC (表 1) に格納し、その映像の鑑賞を行った。

Ladybug は水平方向に 5 個、上方向に 1 個の CCD カメラを配置した小型カメラヘッドと HDD アレイによる蓄積ユニットを用いることで、下方向を除く全天球の約 75% の動画像 (解像度: 3840×1920, フレームレート: 15fps) を撮影することが出来る。今回、鑑賞用 PC の性能の制約から実験に用いた動画像は解像度を 1024×512 に縮小したものをを用いた。また、動画像のフォーマットは、映像部には復号化が高速である理由から MPEG1 を、音声部には MPEG1 Layer2 を用いた。

図 3, 4 に実験時のブラウザ画面を示す。映像と音声は同期して再生され、提示画像は 30fps で更新された。蓄積型動画像の場合、全方位動画像ビューアは動画像を一時停止して周囲を見回したり、早送りをして見たいシーンを探すなどの操作も可能である。

より臨場感の高い鑑賞方法として、図 5 に示すような HMD に利用者の視線方向を検出するための姿勢センサを取り付けた視線追従型の画像提示も行った。今回、姿勢センサには INTERSENSE 社の InterTrax2 を用いた。このセンサは姿勢に関する情報を最大 256Hz で取得可能であり、取り付けやプログラムの実装も容易という特徴がある。姿勢センサ付き HMD を装着し鑑賞を行ったところ、33ms 以下の時間遅延で 360 度視線方向を変えることが出来た。

B) リアルタイム型動画像

リアルタイム型動画像を用いた実験では図 6 に示すように、全方位カメラ HyperOmniVision[6]を搭載した車両、全方位画像を中継しマルチキャストで配信す

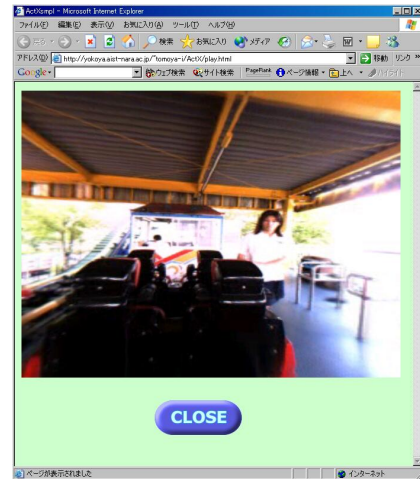


図 3 : 蓄積型動画像使用時のブラウザ画面 1

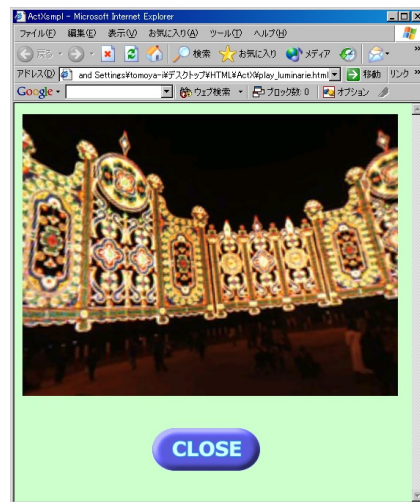


図 4 : 蓄積型動画像使用時のブラウザ画面 2

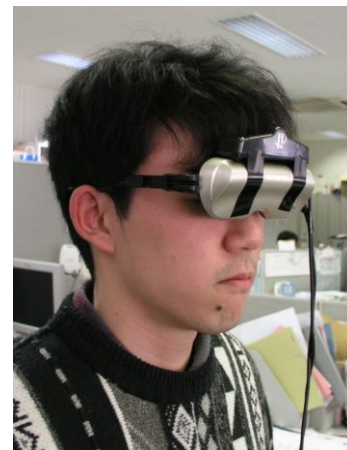


図 5 : ジャイロセンサを取り付けた HMD

るサーバー、全方位動画像ビューア、またそれらをつなぐ無線・有線ネットワークからなるシステムを構築した。全方位カメラを搭載した車両は、本学屋外の道路を走行しながら全方位カメラで周囲 360 度をプログレッシブで撮影し、i.LINK で車内の画像取得圧縮符号

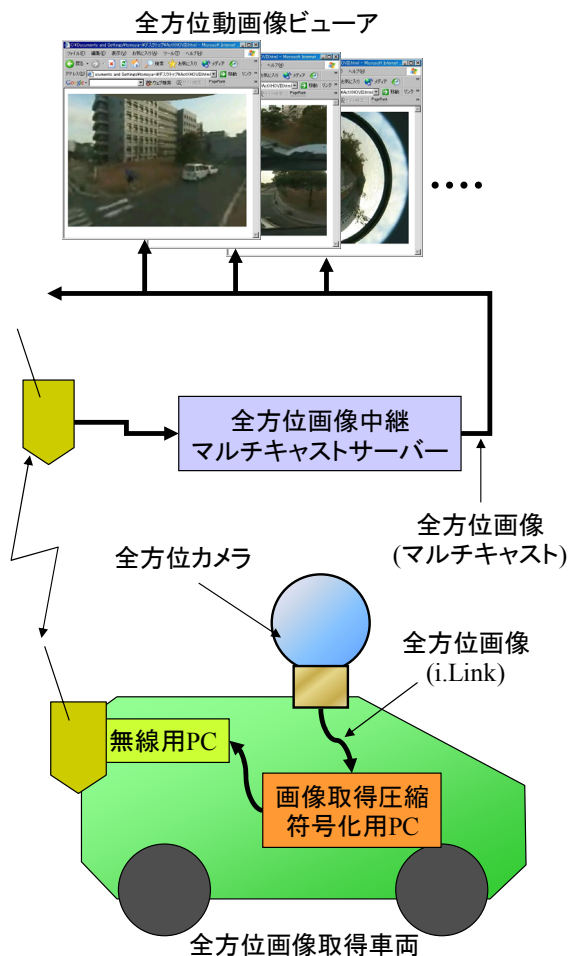


図 6 : 車両全方位カメラからのテレプレゼンス

化用 PC に画像を送る。圧縮符号化用 PC では、全方位動画像（解像度：640×480、フレームレート：30fps）を Windows Media Encoder[7]によって Windows Media 形式（ビットレート：1 Mbps）にエンコードする。エンコードされた動画像は無線用 PC を通じて IEEE802.11a または IEEE802.11g ネットワークにより TCP コネクションなどで本学屋内の中継サーバーに送られる。表 2 に画像取得車両システムの構成を示す。

画像取得車両から送られてきた動画像は全方位画像中継マルチキャストサーバーで受信される。中継サーバーは全方位動画像を RTSP などのプロトコルによりマルチキャストでネットワークに配信する。そして中継サーバーにより配信された動画像を鑑賞用 PC で全方位動画像ビューアを用いて鑑賞する。ユニキャストではなくマルチキャストで配信することにより受信するシステムが多数あったとしてもネットワークの負荷は増加しない。ネットワークがマルチキャストに対応していれば多数の利用者がビューアにより全方位動画像を受信することが可能である。

今回は研究室内の 4 台の PC で全方位動画像を鑑賞

表 2 : 画像取得車両のシステム構成

全方位カメラ	SONY DCR-TRV900 + アコウル 双局面ミラー (視野:水平より上 30 度)
画像取得圧縮符号化用 PC	Pentium4 2.53GHz Memory 1GB WindowsXP
無線ネットワーク	IEEE802.11a (屋外使用免許あり)
車両	日産 ELGRAND (H9 年登録) (図 7, 8)



図 7 : 全方位画像取得車両外観



図 8 : 全方位画像取得車両搭載システム

しようとしたところ、4 台ともそれぞれ任意の方向の画像を見ることが出来た。PC4 台それぞれの全方位動画像ビューアの画面例を図 9, 10, 11, 12 に示す。画像は受信した動画像と同様の 30fps で提示され、カメラの撮影と画像提示の間には約 10 秒の時間遅延があった。時間遅延の主な原因は画像符号化によるものだった。中継サーバーのネットワーク負荷は送受信ともに 1 Mbps であり、画像を受信するシステムの数が増えてもネットワークの負荷が増加しないことを確認した。

4. まとめと今後の展開

本研究では、Webブラウザを用いて全方位動画のインタラクティブ鑑賞が可能な簡易型テレプレゼンスシステムを開発した。そして本システムを用いることにより、蓄積型動画の実験では高品質な動画を利用出来る事を確認した。また、リアルタイム型動画の実験では無線・有線ネットワーク上でマルチキャスト配信を行い、複数の利用者が同時にそれぞれ任意の方向の動画を鑑賞するインタラクティブTV放送のような利用が可能であることを確認した。そして、利用者が容易にコンテンツを鑑賞・操作できることを確認した。

今回リアルタイム型全方位動画の取得に用いた全方位カメラはNTSCの解像度で全方位の画像を取得するため、平面透視投影変換した画像に粗さが目立ち没入感を損なっていた。一方、蓄積型全方位動画の取得に用いたLadybugはマルチカメラシステムという構造上、解像度の高い画像が得られるもののリアルタイムで全方位動画を配信するには適していなかった。そこで今後はHDカメラを用いたHyperOmniVisionを使用することにより高解像度で臨場感の高いネットワークテレプレゼンスを目指す。

また、今後は中継サーバーも車両の中に搭載し、画像取得車両から無線ネットワークを通して直接マルチキャストで画像を配信する実験を行う。

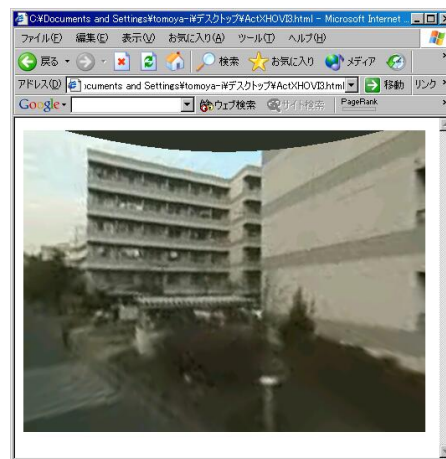
さらに、今回行なったリアルタイム型動画の実験ではカメラの撮影から画像提示の間に約10秒の時間遅延があり、遠隔地の相手とのコミュニケーション等の用途には利用が困難であると考えられる。今後は遅延の少ない映像伝送方法の利用も検討する。

参考文献

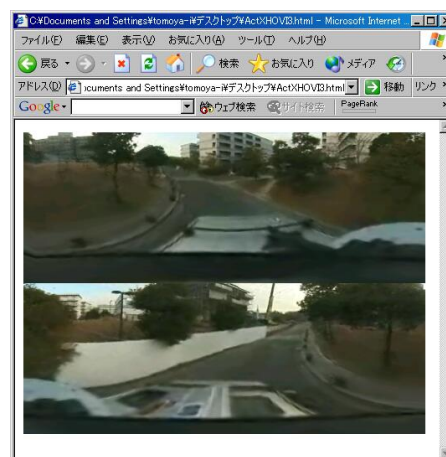
- [1] S.Moezzi, Ed., "Special issue on immersive telepresence", IEEE MultiMedia, vol.4, no.1, pp.17-56, Jan-March 1997.
- [2] Y.Onoe, K.Yamazawa, H.Takemura, and N.Yokoya, "Telepresence by Real-time View-dependent Image Generation from Omnidirectional Video Streams", Computer Vision and Image Understanding, vol.71, no.2, pp.154-165, 1998.
- [3] 池田 聖,佐藤 智和,横矢 直和,"全方位型マルチカメラシステムを用いた高解像度な全天球パノラマ動画の生成とテレプレゼンスへの応用",日本バーチャルリアリティ学会論文誌, vol.8, no.4, pp.443-450, Dec. 2003.
- [4] S.E.Chen,"QuickTime VR - An image-based approach to virtual environment navigation", Proc. SIGGRAPH'95, pp.29-38, 1995.
- [5] 土居 意弘,山本 強,"標準映像フォーマットを用いた実時間全周動画の中継システム",電子情報通信学会 情報・システムソサエティ大会講演論文集, D-11-50, pp.135, 2001.
- [6] 山澤 一誠,八木 康史,谷内田 正彦,"移動ロボットのナビゲーションのための全方位視覚系

HyperOmniVisionの提案", 電子情報通信学会論文誌, vol.J79-D-II, no.5, pp.689-707, May 1996.

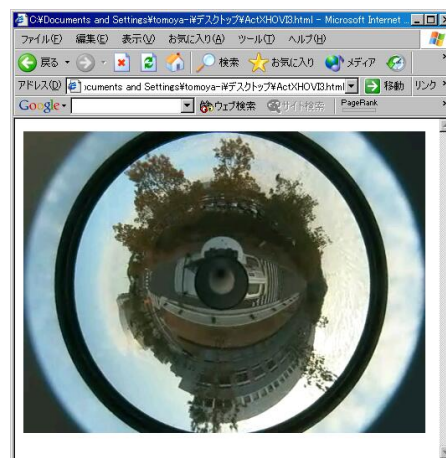
- [7] Microsoft Corporation, Windows Media Encoder 9 Series, <http://www.microsoft.com/windows/windowsmedia/9series/encoder/default.aspx>



平面透視投影表示



パノラマ(円筒画像)表示



全方位画像(入力画像)表示

図9：全方位動画ビューアの画面例 (ユーザー1)



図 1 0 : 全方位動画像ビューアの画面例 (ユーザー 2)

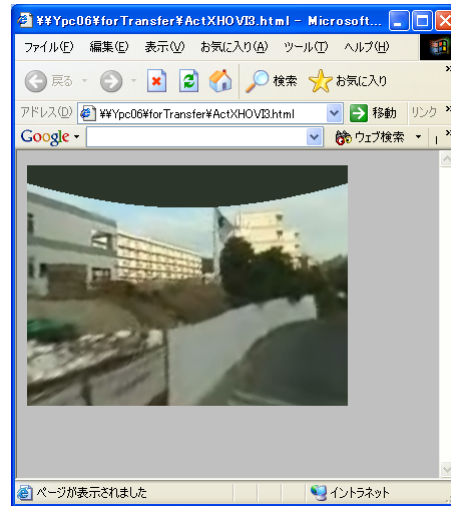
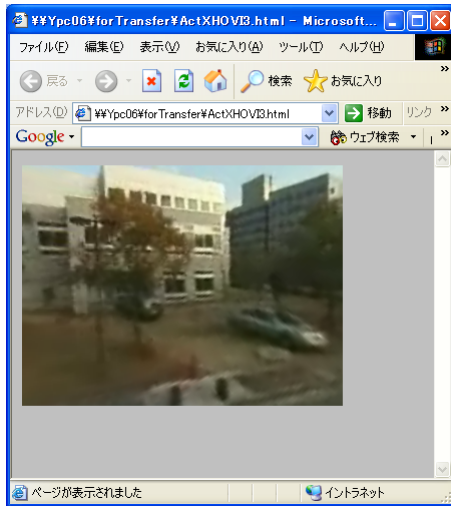


図 1 1 : 全方位動画像ビューアの画面例 (ユーザー 3)

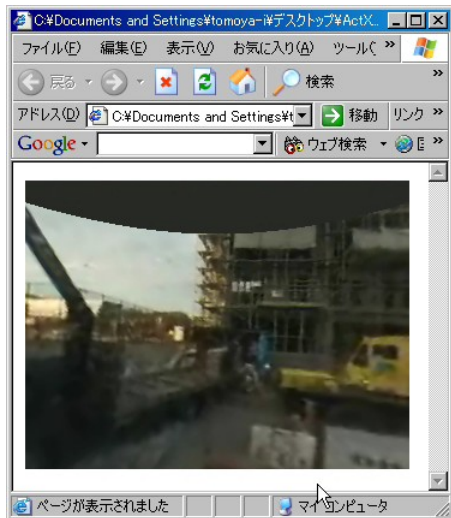


図 1 2 : 全方位動画像ビューアの画面例 (ユーザー 4)