

全方位ビデオカメラを用いた視覚情報メディア

横 矢 直 和[†] 山 澤 一 誠[†] 竹 村 治 雄^{†,‡}

近年、レンズ・ミラー系の組合せによって側方 360 度の視野をビデオレートで撮影できる様々な全方位画像センサが開発されるようになってきた。このような全方位ビデオカメラの中でも特に、1 点中心投影の光学特性を有するカメラは、撮影された全方位画像から任意形状のディスプレイ・スクリーン面への透視投影画像を計算によって生成できることから、最終的に人間への画像・映像提示を目的としたメディア応用に適している。本論文では、全方位画像からの任意視線画像の実時間生成による、視覚情報メディアへの全方位ビデオカメラの応用について述べる。具体的には、全方位ビデオカメラ HyperOmni Vision を用いて筆者らの研究室で進めてきたテレプレゼンス、代理身体としての移動ロボットの遠隔操縦、およびビデオサーベイランスに関する一連のプロジェクトの概要について述べる。

Visual Information Media Using Omnidirectional Video Cameras

NAOKAZU YOKOYA,[†] KAZUMASA YAMAZAWA[†]
and HARUO TAKEMURA^{†,‡}

Recently an increasing number of catadioptric video-rate omnidirectional imaging sensors with 360-degree of horizontal view have been developed. Among those omnidirectional video cameras, ones which satisfy the single-viewpoint constraint are suitable for emerging interactive media-oriented applications, because perspective re-projection onto any display or screen surfaces can be computed from captured omnidirectional images. This omnibus paper describes a number of applications of catadioptric omnidirectional video camera HyperOmni Vision to visual information media, especially focussing on our recent activities including telepresence, teleoperation of mobile robot, and video surveillance. All of these applications are based on computing view-dependent perspective images from omnidirectional video streams in real time.

1. はじめに

全周パノラマビュー・カメラのアイデア¹⁾は 1960 年代にまでさかのぼることができるが、近年、レンズ・ミラーの組合せによる反射光学系や複数のビデオカメラを用いることによって、動的な環境を対象として、側方 360 度の視野をビデオレートで撮影できる様々な全方位画像センサ(以下、全方位ビデオカメラとも呼ぶ)が開発されるようになってきた^{2)~4)}。また、これらを用いたロボット視覚等、様々な応用事例も報告されている²⁾。

全方位ビデオカメラの中でも特に、1 点への中心投影の光学特性を有し単一視点制約を満たす全方位ビデオカメラは、撮影された全方位画像から任意方向・任

意形状のディスプレイ・スクリーン面への透視投影画像を計算によって生成できることから、最終的に人間への画像・映像提示を目的とした視覚情報メディアへの応用に適している^{5),6)}。また、最近の仮想現実・複合現実⁷⁾応用では、実環境を対象とした没入型の映像空間を構築するために、実環境を撮影した全方位ステレオ画像^{8),9)}や多視点全方位画像からの自由視点・視線画像生成^{10),11)}も行われるようになってきた。これらの研究は仮想化現実とも呼ばれる。遠隔世界にいたかにも実際にいるかのような没入感を与える技術にテレプレゼンス¹²⁾があるが、没入感が得られるための大きな要因の 1 つは、観察視点・視線の自由な変更に従った時間遅れの少ない画像の実時間提示である。

本論文では、単一視点制約を満たす全方位画像からの任意視線画像の実時間生成による仮想カメラのパン・チルト・ロール操作を利用した、視覚情報メディアへの全方位ビデオカメラの応用事例を紹介する。具体的には、全方位ビデオカメラ HyperOmni Vision¹³⁾を用いたインタラクティブ画像・映像提示の応用として

[†] 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology
現在、大阪大学サイバーメディアセンター
Presently with Cybermedia Center, Osaka University

筆者らの研究室でこれまでに行ってきたテレプレゼンス、代理身体としての移動ロボットの遠隔操縦、およびビデオサーベイランスの一連のプロジェクトに関して、システムの実装例の紹介に重点を置いて述べる。

以下、まず 2 章では、本研究で用いる全方位ビデオカメラの概要と全方位画像からの提示画像の生成法について述べる。3 章では、全方位画像からの提示画像の実時間生成の具体的な応用として、いくつかのテレプレゼンスシステムの構成例を示す。4 章では、代理身体による遠隔鑑賞・観察を目的とした、テレプレゼンス機能を用いた移動ロボットの遠隔操縦について述べる。5 章では、複数の全方位ビデオカメラを用いたビデオサーベイランスシステムについて述べる。最後に 6 章で一連の研究のまとめを行う。

2. 全方位画像からの提示画像の生成

2.1 全方位ビデオカメラ

メディア応用では、最終的には人間への画像提示を行うため、人間への提示に適した画像を全方位画像から生成する必要がある。その場合、ディスプレイ面に映っている画像の網膜への投影が、通常の実シーンの投影と同じであることが望ましい。このためには、網膜への投影画像が近似的に平面透視投影に従っている必要がある。

ディスプレイ・スクリーン面の形状は、通常の計算機のモニタ、頭部搭載型画像表示装置 (HMD: Head-Mounted Display) 等における平面、CAVE 型スクリーンにおける直方体表面、アーチ型スクリーンにおける円筒面、ドーム型スクリーンにおける球面等、様々である。これらに透視投影画像を表示するためには、元の全方位画像が形状が既知の投影面への 1 点中心投影である必要がある。この制約は単一視点制約と呼ばれる¹⁴⁾。レンズ・ミラー系の組合せによる全方位画像センサにおいて単一視点制約を満たすミラー形状は平面¹⁵⁾、双曲面¹³⁾、放物面¹⁶⁾等に限定されることが知られている。

本研究では、単一視点制約を満たす全方位ビデオカメラの 1 つである HyperOmni Vision¹³⁾を用いる。同センサは、図 1 に示すように、鉛直下向きに設置した回転 2 葉双曲面ミラーの回転軸上の外側焦点位置 (O_C) に投影中心が一致するようにビデオカメラを鉛直上向きに設置し、側方 360 度の情景の双曲面での鏡像を撮像する構成になっている。下方視野はカメラ自体が映り込む部分を除いて 90 度に近く、上方視野は双曲面ミラーの形状により変化する。

図 1 の反射光学系において、外界から双曲面ミラー

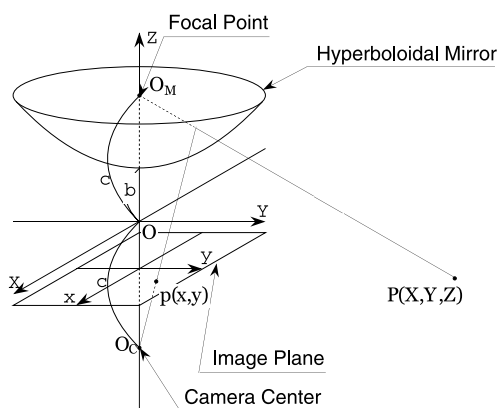


図 1 全方位ビデオカメラ HyperOmni Vision の光学系
Fig. 1 Geometry of HyperOmni Vision.

の内側焦点 O_M に向かう光線はミラーで反射されて外側焦点 O_C を通過する。ここで、図 1 に示すように、 Z 軸を鉛直軸とする 3 次元世界座標系 (X, Y, Z) と画像座標系 (x, y) を考えると、カメラがピンホールカメラモデルに従う場合、3 次元空間中の点 $P(X, Y, Z)$ と全方位画像中の写像点 $p(x, y)$ の間には次式の関係が成り立つ。

$$\begin{aligned} x &= \frac{f(b^2 - c^2)X}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}, \\ y &= \frac{f(b^2 - c^2)Y}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}. \end{aligned} \quad (1)$$

ただし、 f はカメラの焦点距離であり、 a, b, c は双曲面ミラーの形状を表すパラメータである。このように、HyperOmni Vision は単一視点制約を満たし、内側焦点 (O_M) を投影中心とした双曲面への 1 点中心投影像が得られる。

2.2 透視投影画像の生成

HyperOmni Vision は上で述べたように、双曲面ミラーの内側焦点位置を投影中心とする単一視点制約を満たすため、式 (1) をもとに、取得した全方位画像から任意の位置・向きに設定された任意形状の投影面への再投影画像を計算によって求めることができる。取得した全方位画像から平面透視投影画像を生成した例を図 2 に示す。また、円筒面への再投影によって求めた全周パノラマ画像を図 3 に示す。これらの透視投影画像は、基本的に、全方位画像の幾何学的変換 (座標変換) によって得られる。

全方位画像からの平面透視投影画像の生成により、通常のカメラの視点固定でのパン・チルト・ロール・ズーム操作に対応した仮想カメラ操作を実現すること

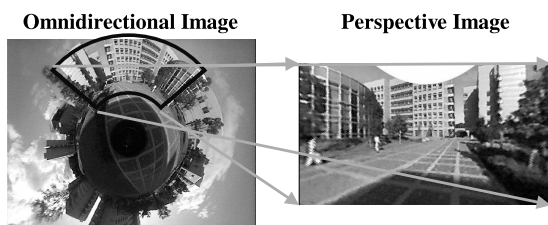


図2 全方位画像からの平面透視投影画像の生成

Fig. 2 Computing a common perspective image from an omnidirectional image.



図3 全方位画像からの全周パノラマ画像（円筒面透視投影画像）の生成

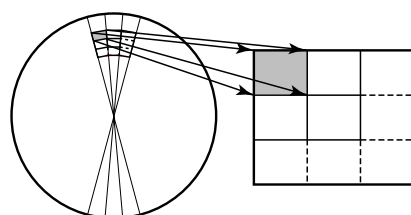
Fig. 3 A full panoramic image computed from an omnidirectional image.

ができる．実カメラを用いたこのような操作では一般に，機械的な動作遅延が発生し，また，カメラが遠隔地にある場合には通信遅延や画像の伝送遅延も無視できなくなる．全方位画像からの視線に追従した平面透視投影画像生成を観察者側で実時間で行うことができれば，実カメラのような時間遅れの問題なく，遠隔地の情景を自由に見回すことができる．また，この場合には，実カメラを用いる場合と違って，仮想的なカメラ操作を同時に複数人が独立に行うことができるという利点もある．

2.3 画像生成の高速化

動環境のインタラクティブな観察を実現するためには，全方位ビデオストリームから前節で示したような提示画像をユーザの視線に追従して実時間で計算する必要がある．現状のワークステーション等でソフトウェア的にこれを実現するのは難しい．そこで，変換後の画像内の比較的少数の格子点上でのみ対応する入力全方位画像内の座標を計算し，格子間の隙間は画像変形（イメージワーピング）¹⁷⁾のためのハードウェア機能を利用して補間する疑似透視投影画像生成法（図4参照）を開発している¹⁸⁾．本手法では厳密には透視投影画像を生成できないため，生成画像には幾何学的な歪みが存在する．このため，人間が不自然さを感じない，すなわち，幾何学的な歪みを知覚できない程度の画像を生成できる格子点数を設定する必要がある．

本手法は，任意形状の投影面に対応可能であり，平面透視投影画像の生成以外に，全方位画像から円筒



Source Image Destination Image

図4 画像変形による全方位画像からの平面透視投影画像の生成

Fig. 4 Image warping for generating a planar perspective image from an omnidirectional image.

面への再投影画像である全周パノラマ画像（図3）やCAVE型スクリーンへの再投影画像¹⁹⁾の実時間生成にも適用できる．

3. テレプレゼンス

時間的あるいは空間的に離れた遠隔実世界にあたかも実際にいるかのような没入感を与える技術にテレプレゼンスがある．没入感が得られるための大きな要因の1つは視点・視線の自由な変更による観察のインタラクティブ性である．前章で述べた全方位画像からの透視投影画像生成方式を用いて視線の自由な変更による見回し機能を実現することによって，以下のようなテレプレゼンスシステムを構築できる²¹⁾．いずれもライブビデオによる実時間型とテープ等への記録ビデオを用いた蓄積再生型での利用が可能である．

- (1) スタンドアロン型システム：
入力した全方位ビデオストリームから視線に追従した透視投影画像を実時間で生成・提示する．入力ビデオ信号を分岐すれば複数人の独立視聴も可能になる．
- (2) ネットワーク型システム：
1つのサイトからネットワークを介して多地点に全方位ビデオストリームをマルチキャストし，受信側で(1)と同様の処理を行う．本方式では，1つのビデオストリームをもとに複数人が異なる視聴を行うことができる．
- (3) 放送型システム：
放送網を利用して全方位ビデオストリームをブロードキャストすることによって，(2)と同様の機能を実現する．

これらのシステムでは，視線情報の与え方によって様々なユーザインタフェースが考えられる．たとえば，

テレイグジスタンスあるいは remote reality²⁰⁾とも呼ばれる．

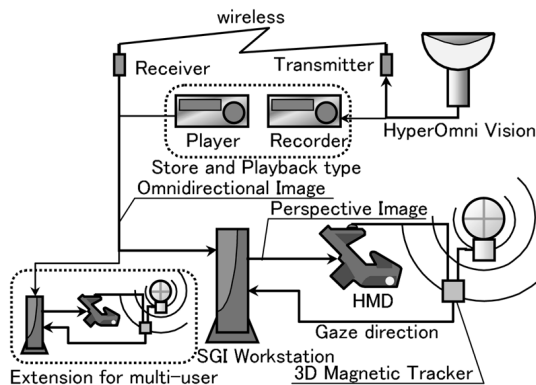


図5 HMDを用いたテレプレゼンスシステムの構成

Fig. 5 Configuration of a telepresence system using HMD.

視線検出，頭部追跡，ジョイスティックやマウスの利用等がある．以下では，全方位ビデオストリームのインタラクティブな観察・観賞を可能とする3種類のスタンドアロン型プロトタイプシステムについて述べる．

3.1 HMDを用いたテレプレゼンスシステム

HyperOmni Vision と3次元磁気トラッカ付きHMDを用いた実時間/蓄積再生型テレプレゼンスシステムの構成を図5に示す^{(18), (22)}．本システムでは，ユーザの視線情報は頭の向きを変えるという自然な動作で与えられる．この場合，ユーザの視線変化に追従した実時間画像提示が不可欠であり，さらに視線の変更から画像提示までの時間遅延を極力抑える必要がある．従来の可動カメラを用いたシステムでは，ユーザの視線方向に応じてカメラを機械的に動かすための時間遅延が生じ，また，遠隔地の観察では通信時間による遅延も無視できないが，全方位ビデオカメラを用いることによってこれらの問題点を解決するとともに，複数人の同時独立使用が可能なマルチユーザシステムの構成が可能になる．

プロトタイプシステムで使用した機器の詳細を表1に示す．実時間型システムにおける全体の処理の流れは以下のとおりである．

- (1) HyperOmni Vision で撮影した全方位画像を計算機に無線伝送する．
- (2) ユーザの装着するHMDに取り付けた3次元磁気トラッカでユーザの視線（頭部）の方向（3自由度）を計測する．
- (3) 計算機でユーザの視線に対応した平面透視投影画像を生成し，HMDに表示する．

蓄積再生型システムでは，ステップ（1）において記録された全方位ビデオを再生し，ステップ（2）と（3）

表1 テレプレゼンスシステムの使用機器

Table 1 Equipments used in telepresence system.

全方位画像センサ	HyperOmni Vision ver.2A
画像無線送受信器	Premier Wireless CS-200
計算機	SGI Indigo2 (R4400, 250 MHz) Maximum Impact (Texture Engine: 119 Mpixels/s) Impact Video
HMD	OLYMPUS Mediamask (512, 880 pixels) (Input: S-Video NTSC)
磁気トラッカ	POLHEMUS 3SPACE FASTRAK

は実時間型システムと同じである．

ステップ（3）での画像生成には2.3節で述べた高速化手法を用いている．この場合，先にも述べたように，格子点が少ないと，画像の幾何学的な歪みが目立つため，生成画像に歪みが目立たない程度の格子点数を設定する必要がある．画像生成時間と画像の幾何学的歪みに関する評価実験を行い，全方位画像，生成画像ともに 720×486 画素の場合，表1の機器構成においては，ビデオレートの画像更新を実現するために画像生成時間を $1/30$ 秒以下に抑えるためには格子数は 32×24 以下，幾何学的歪みが視覚的に目立たないためには格子数は 16×12 以上という結論を得ている⁽²²⁾．これらの値は計算機の性能とHMDの解像度に依存する．

本システムでのユーザの視線変化から対応する画像表示までの時間遅延は，ビデオ映像の実時間伝送が可能である限り，遠隔地との通信時間や画像伝送時間には依存せず，視線検出および画像生成・提示に要する時間で決まる．現状のシステム構成における時間遅延は $51.5 \sim 93.1$ ミリ秒である．この程度の時間遅れであれば，ユーザは頭を速く動かした場合に提示画像の時間遅れを知覚するが，ほとんど違和感を感じることなく見回しが行えることを実験により確認した⁽²²⁾．ただし，ユーザは通信時間と画像転送時間に応じた過去の情景を見ていることになる．

本システムでは，ユーザの視線情報としてユーザの頭部の3軸回りの回転角度を計測しているため，仮想カメラのパン・チルト・ロール操作の結果がHMDに提示される．ただし，観察視点はHyperOmni Visionを構成する双曲面ミラーの内側焦点位置（図1中の点 O_M ）に固定されており，観察視点の移動は撮影時のセンサの移動によってのみ可能である．

開発したプロトタイプシステムを用いて，谷瀬の吊り橋（奈良県吉野郡十津川村）を歩行しながらHyperOmni Visionで撮影した全方位ビデオストリームをも

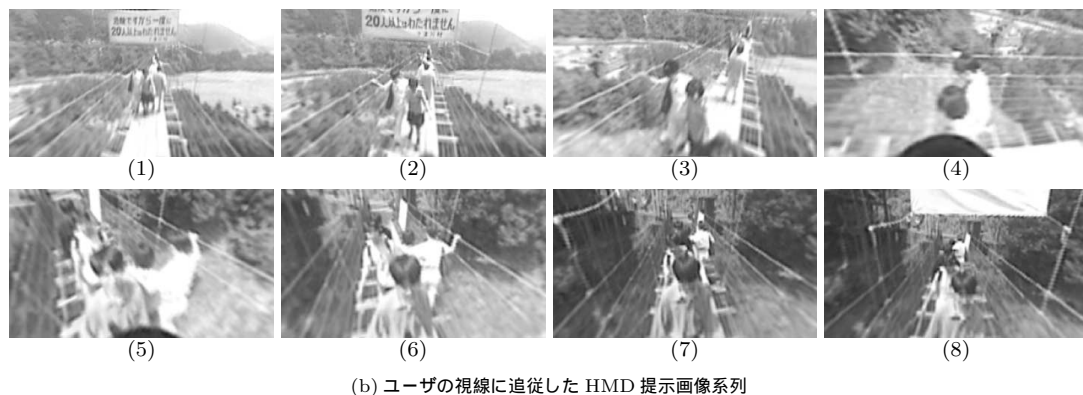
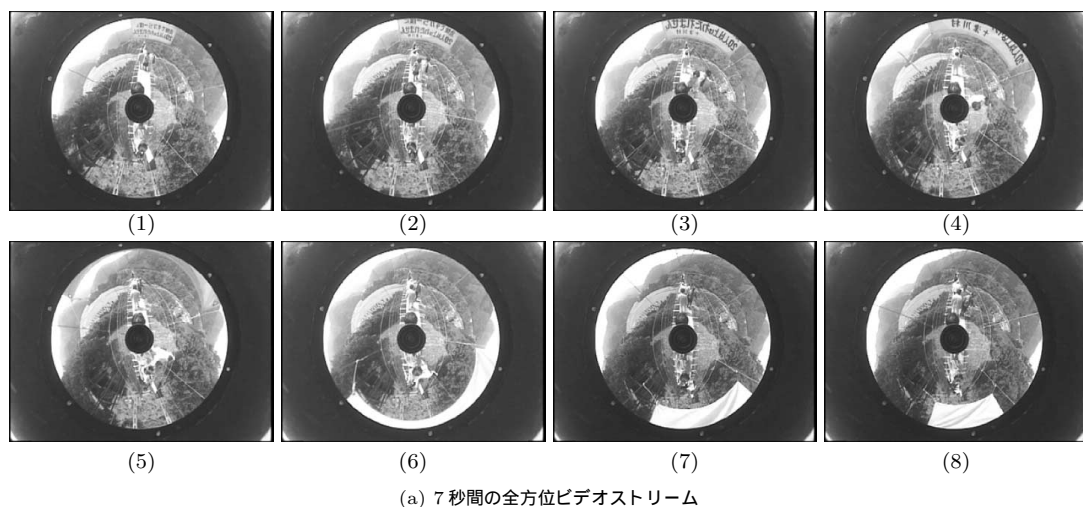


図 6 HMD を用いたテレプレゼンスシステムによる実環境（谷瀬の吊り橋）の仮想体験
Fig. 6 Virtual tour into a virtualized dynamic real environment (Tanise Suspension Bridge) using HMD-based telepresence system.

とに，上記のステップ（1）であらかじめ撮影された全方位ビデオを再生しながらの蓄積再生型テレプレゼンスを実現した場合の全方位画像と HMD 提示画像の系列を図 6 に示す．

3.2 IPD を用いたテレプレゼンスシステム

単一視点制約を満たす全方位画像は，任意形状のスクリーン面を持つ没入型投影ディスプレイ（IPD: Immersive Projective Display）への対応も容易である．ここでは，没入型投影ディスプレイでの実現例として，入口を除く側方 330 度の視野を有する大型円筒スクリーン（直径 6 m，高さ 2.4 m）とグラフィックスワークステーション等からなる全周型景観提示システム CYLINDRA を用いたテレプレゼンスシステムを紹介する．

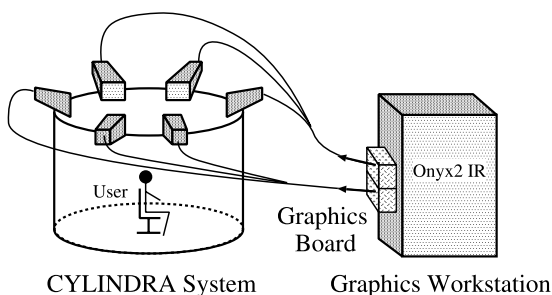


図 7 全周型景観提示システム CYLINDRA の構成
Fig. 7 Configuration of an immersive projective display CYLINDRA.

CYLINDRA のシステム構成を図 7 に示す．本装置は 330 度の視野を 6 台のプロジェクタで分割表示する構成になっており，各プロジェクタへの画像を全方位画像から実時間で生成・送信する．なお，現状の構成では 1 つのグラフィックスボードが 3 つのプロジェ

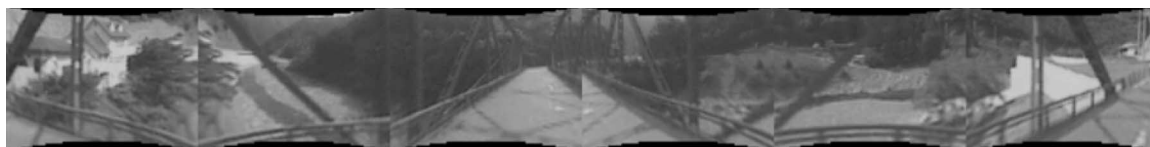


図 8 CYLINDRA に投影する 6 枚の円筒面画像

Fig. 8 Six cylindrical images projected onto a screen of CYLINDRA.



図 9 CYLINDRA での実環境の仮想体験

Fig. 9 Virtual tour into a virtualized dynamic real environment using CYLINDRA.

クタへの画像提示を担っている。

本装置は全周スクリーンを有するため、前節のHMDを用いた実現例のように観察者の視線に追従した画像を生成する必要はなく、つねに円筒スクリーンの中心を視点とした図3のような全周パノラマ画像を計算すればよい。ただし、円筒面への内側からの映像投影を考慮した透視投影画像の生成が必要である。提示画像の実時間生成には、2.3節で述べた方法と同様に画像変形を用い、HyperOmni Visionで取得した全方位画像から全周パノラマ画像を生成するのに 96×12 の格子（6枚の提示画像それぞれについて 16×12 ）を用いてビデオレートでの画像計算を実現している。6台のプロジェクタからの投影画像の例を図8に、また、車載 HyperOmni Visionで取得した全方位ビデオを用いた、CYLINDRAでの実環境の仮想体験の様子を図9に示す。

ここでは、円筒面スクリーンからなるIPDでの実装例を紹介したが、CAVEに代表されるキューブ型や球面ドーム型のスクリーンからなるシステムへの実装も容易である。実際に、キューブ型4面スクリーン（前面、下面、左右面）においても同様のシステムが構築されている¹⁹⁾。

3.3 ステレオテレプレゼンスシステム

前節までに紹介したテレプレゼンスシステムにおいては、ユーザは観察視線を自由に選ぶことができるが、

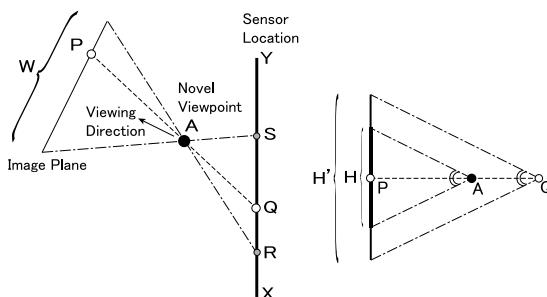


図 10 全方位画像系列からの任意視点画像の生成

Fig. 10 Illustration of generating a novel view from a sequence of omnidirectional images.

観察視点は全方位ビデオカメラを構成する双曲面ミラーの内側焦点位置に固定されており、自由視点画像や両眼ステレオ画像の提示はできない。テレプレゼンスでは、遠隔地の情景を臨場感豊かに提示することが重要であり、時間遅延のない見回しに加えて、立体視が可能であることが望まれる。

近年、蓄積された画像群から自由視点画像を生成するイメージバーストレンダリングの一手法として、3次元空間を伝播するすべての光線を位置、方向、波長、時間の関数として記述する plenoptic function²³⁾を利用した描画手法がいくつか提案されている（たとえば、文献24）、25）。本節では、撮影位置が既知である全方位動画から光線情報を利用して自由視点画像を疑似的に生成する手法により、視線に追従した両眼ステレオ画像の提示を可能とするテレプレゼンスシステム¹⁰⁾について述べる。

ここでは、HyperOmni Visionを直線移動させながら全方位動画を取得することを考える。図10に示すように、直線XY上をセンサが移動して経路上で全方位画像を取得すると、経路上以外の新しい視点Aでの平面透視投影画像（大きさ $W \times H$ ）は、同図に示すように、区間RS上で取得した全方位画像から疑似的に生成することができる。たとえば、生成画像面上の画素Pは経路上の点Qで撮影された全方位画像から計算される。ただし、本手法では、生成画像の縦方向に全方位画像の撮影位置から物体までの距離に依存した歪みが生じる（縦方向の歪みに関する解析につい

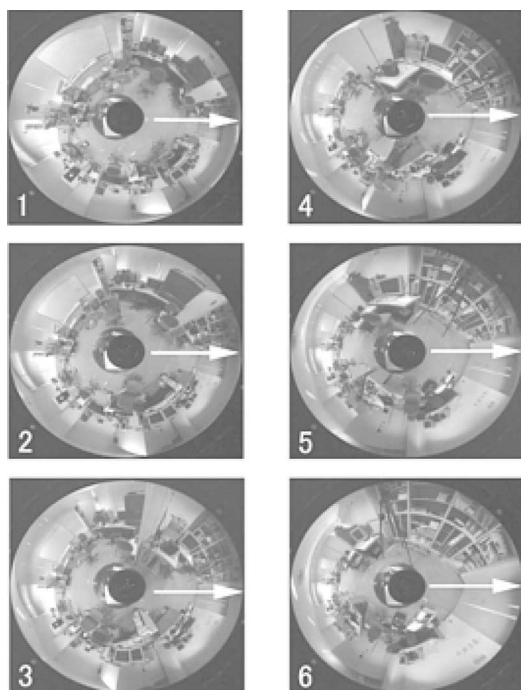


図 11 入力全方位動画画像

Fig. 11 A sampled input sequence of omnidirectional images.

ては文献 26) 参照)。

この疑似的な自由視点画像生成法を用いて、両眼中心をセンサの移動経路上に設定した 2 視点での画像を生成することによって両眼立体視の可能なテレプレゼンスシステムを構築した。システムは図 5 と似た構成であり、ある一定時間の全方位画像が計算機に蓄積される。実験では、HyperOmni Vision を移動ロボット (Nomad-200) に搭載し、17.5 cm/s の一定速度で直進させ、計算機にはつねに最新の 2.03 秒間の全方位動画画像を保持しながら、その動画画像の撮影中間位置における両眼ステレオ画像を生成し HMD に提示した。すなわち、計算機に保持されている全方位画像は最新の 61 枚であり、両眼ステレオ画像を生成する地点から前後 30 枚の全方位画像がつねに計算機に蓄積されている。そのため、ユーザには約 1 秒前の地点での両眼ステレオ画像が提示される。なお、本システムで生成される両眼ステレオ画像の眼間距離は 7 cm、水平画角は 50 度、解像度は 640×480 画素である。

図 11 に、実験室内で取得した約 35 秒間の入力全方位動画画像中の 7 秒間隔の全方位画像を示す。なお、同図中の矢印はロボットの進行方向を表している。図 12 はユーザが見回しを行った場合に HMD に提示された両眼ステレオ画像である。本手法には、視線方向が

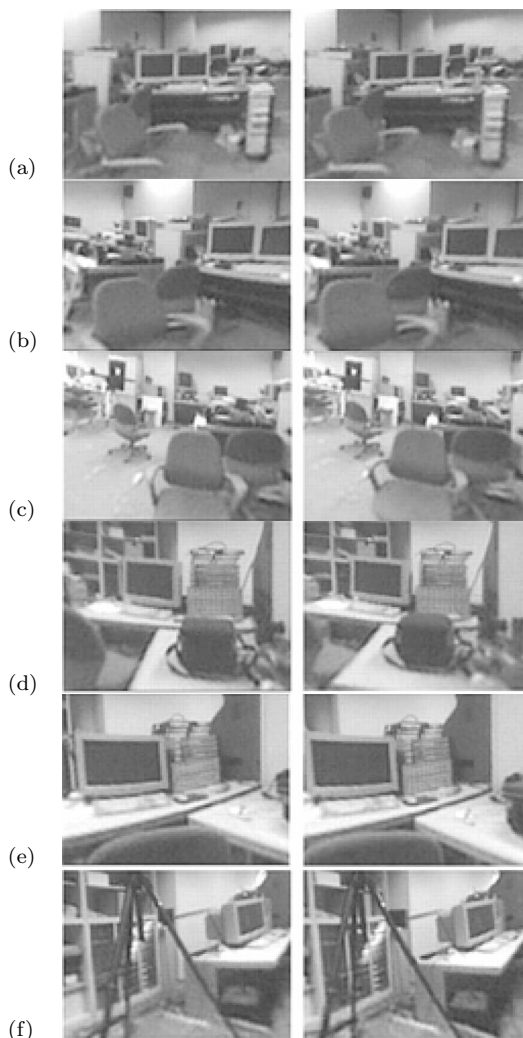


図 12 生成された両眼ステレオ画像

Fig. 12 A sequence of generated binocular stereo images.

ロボットの進行方向に近い場合には提示画像の生成に多数の全方位画像を必要とし、両者が完全に一致する場合には両眼ステレオ画像は生成できないという問題がある。本システムの実装では、両眼ステレオ画像が提示されるのはロボットの進行方向に対して左右 36 度～144 度の範囲であり、それ以外の方向については 3.1 節で述べた手法を用いて、ユーザの両眼中心位置で取得された 1 枚の全方位画像から単眼画像を生成・提示している。視線決定から HMD への画像提示までの時間遅延は 37～70 ミリ秒であり、画像更新はビデオレート (30 枚/秒) であった。また、両眼ステレオ画像生成における縦方向視差 (歪み) は最大 15 画素程度、ほとんどが 5 画素以下であり、融像に影響を与えるほどではなく、ユーザは両眼立体視による奥行き

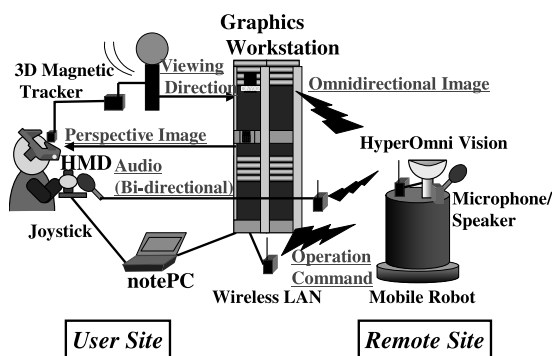


図 13 移動ロボット遠隔操縦システムの構成

Fig. 13 Configuration of a telerobotics system.

知覚が可能であることが確認された²⁶⁾。

4. 移動ロボットの遠隔操縦

3.1 節で紹介したテレプレゼンス機能を用いて、図 13 のような、代理身体による遠隔観賞・監視を目的とした移動ロボットの遠隔操縦システムを開発している²⁷⁾。本システムでは、前進・後退および回転が可能なロボット (Nomad-200) に搭載された HyperOmni Vision から無線伝送される全方位ビデオストリームから視線に追従した透視投影画像を実時間で生成し、HMD に提示することによって、操作者はロボットに乗った感覚でジョイスティックを用いて操縦を行うことができる。また、ロボットに搭載されたマイク・スピーカにより、操作者は遠隔地にいる人物とのコミュニケーションも可能である。

ユーザの操作状態には移動モードと注視モードがあり、それぞれの操作モードでのユーザに対する画像提示は以下になっている。

- 移動モード：ロボットの移動中は、3.1 節で述べたシステムと同様に、HMD を装着したユーザの頭の向き (視線) に追従した単眼画像が HMD に実時間提示される。この場合、ユーザはロボットの移動にともなう運動視差により奥行き知覚が可能である。
- 注視モード：ロボットが停止し、回転しながら取得した全方位画像系列からユーザの視線に追従した両眼ステレオ画像が生成・提示される。なお、ここではロボットの回転軸と全方位センサの軸をずらすことによって両眼ステレオ画像の生成を可能にしている。このモードでは、ユーザは停止した状態で両眼立体視により奥行き知覚が可能である。このように、ユーザは、ロボットの移動時には運動立体視により奥行きの手がかりを得て、停止時には両

眼立体視により環境の奥行きを知覚しながら環境の自由な見回しを行うことができる。

遠隔情景の自由な見回しが可能な映像提示によって、ユーザはあたかもロボットに乗っている感覚で操縦を行うことができるが、見回しによってロボットの移動方向を見失う可能性がある。また、ロボット自身の大きさを把握し周囲の物体までの距離を考慮しながら操縦するためには、頻繁に下方を眺めてロボット自身を見る必要があるという問題もある。このような操作上のユーザインタフェースを改善するために、以下のような操作支援機能を付加した。

(1) 進行方向の重畳表示

全方位画像による自由な見回しが可能な環境では、提示画像中にロボットの一部分が含まれていないと、ロボットの進行方向の把握が困難になるが、今回の実装では円筒形のロボットを用いているため、ロボットの一部分を眺めたとしても、方向の把握は難しい。そこで、ジョイスティック操作における前方をロボットの進行方向に一致させるとともに、ロボットの進行方向を明示的に示す方法として、HMD への提示画像に進行方向を表す矢印の CG を重畳表示している。

(2) 移動経路幅の重畳表示

ロボット自身を眺めることなく、その大きさを把握する方法として、床面上の位置にロボットの足元から進行方向に、ロボットの幅を表す 2 本の直線の CG を重畳表示している。

図 14 に、実験時の様子 (操作者、全方位画像、移動ロボット、HMD 提示画像) を示す。同図中の (e)、(f) は注視モードであり、両眼ステレオ画像が提示されている。それ以外は、移動モードで単眼画像が提示されている。

本システムは、代理身体を用いた遠隔観賞や次章で述べるサーベイランスシステムにおける能動的サーベイランスへの応用が期待できる。また、本論文で述べた走行型ロボットを用いた実装以外に、遠隔操縦型の小型飛行機、ヘリコプタ、飛行船等の飛行体を用いたシステム構成が考えられ、これによって、より広範な応用が可能となる。

マイクとスピーカを利用した遠隔地とのコミュニケーションに関しては、ユーザは遠隔地の情景を自由に見回すことにより遠隔地の人物を視認できるが、逆に、遠隔地にいる人物はユーザをまったく視認できないため、相互認識の低下がコミュニケーションに支障をきたすことがあることが実験で確認された。

本論文の実験では、ユーザのいる場所とロボットの

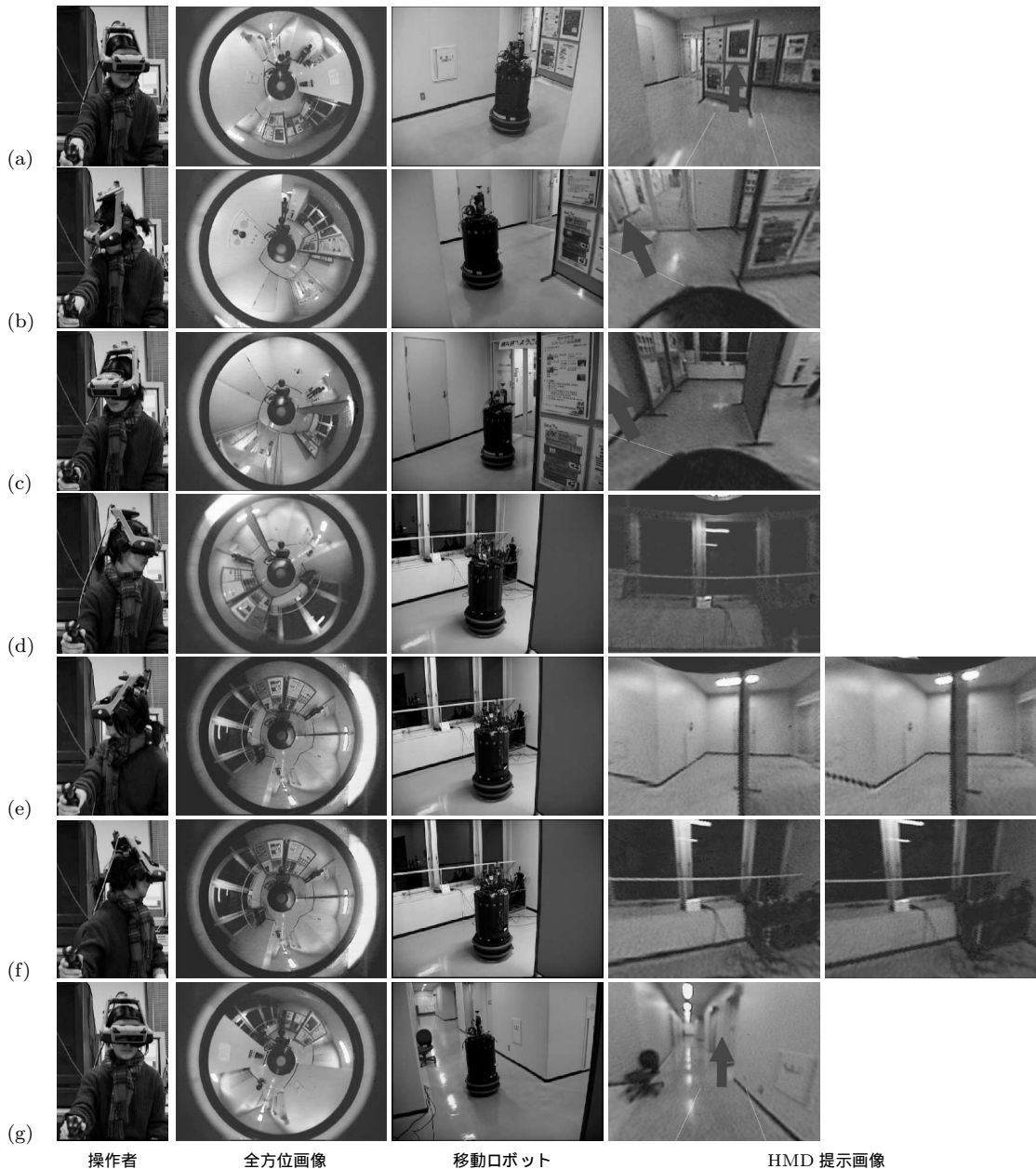


図 14 移動ロボットの遠隔操縦の様子

Fig. 14 A sequence of teleoperating a mobile robot.

存在する場所が比較的近い場合、両地点間での通信時間と画像伝送時間は無視できるが、3.1 節の応用のように遠隔地を眺めるだけでなく、遠隔地での操作をともし本応用では、遠距離になるとこれらの時間遅延を無視できなくなる。そのため、通信時間と画像伝送時間による時間遅延を補うための操作支援手法の開発が必要である。

5. ビデオサーベイランス

広範囲の環境をつねに監視し続けるサーベイランスには、側方 360 度の視野を有する全方位ビデオカメラが適している²⁸⁾。サーベイランスにおける画像提示では、観察者の視線方向に追従した画像提示と移動物体等の変化領域に注目した画像提示が必要であり、いずれも表示遅延を極力抑えた実時間画像生成・提示が不

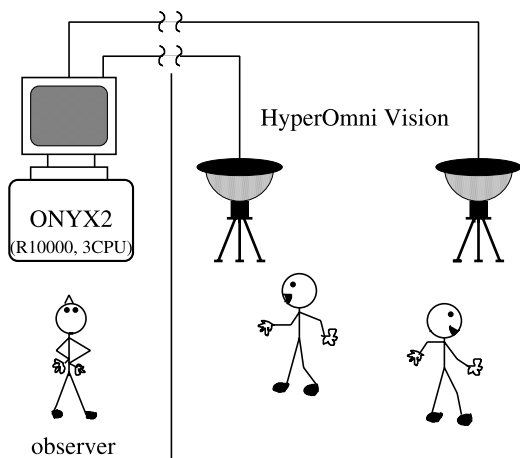


図 15 2 台の HyperOmni Vision を用いたサーベイランス環境
Fig. 15 VSAM using two HyperOmni Visions.

可欠である。

床面上の同じ高さに 2 台の HyperOmni Vision を設置した図 15 のような環境設定で、3.1 節で述べた機能を用いた観察者の意志に基づく遠隔監視に加えて、以下のような機能を有するビデオサーベイランスシステムのプロトタイプを開発している²⁹⁾。

- (1) 全方位画像の背景差分と色情報の利用による移動物体の自動抽出・追跡
- (2) センサ位置から移動物体を見た平面透視投影画像の実時間生成による仮想カメラのパン、チルト操作を用いた自動モニタリング
- (3) 2 視点からの全方位ステレオ視による移動物体の床面上の位置推定

上記 (2) における全方位画像からの平面透視投影画像生成では、2.3 節で述べた画像変形による高速画像生成法を用いており、これら 3 つの機能を備えた実時間 (5 フレーム/秒) サーベイランスが可能である。(1) と (2) により観察者への注意喚起と移動物体に追従した映像提示が行われ、(3) により環境内での物体の移動状態が視覚的に提示される。

図 16 に、実験室内を移動している 2 人の人物の位置推定結果を示す。現状では、2 つのセンサを結んだ線 (ベースライン) 上に存在する物体については原理的に 3 次元位置を特定することができない。このような全方位ステレオの死角を解消するためには 3 台以上のセンサを用いた多視点全方位ステレオ視³⁰⁾が必要である。

6. おわりに

近年、全方位画像センサを用いた研究が活発に行わ

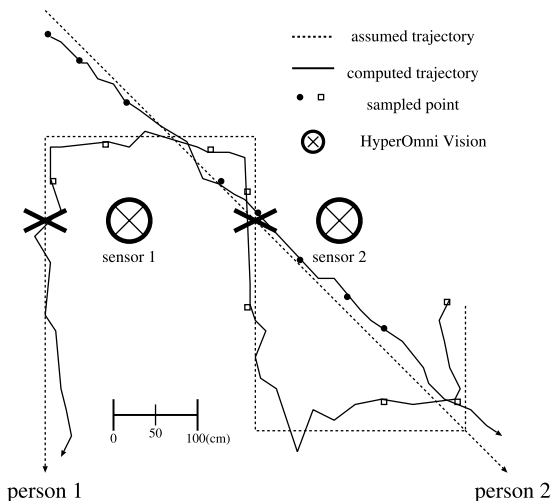


図 16 全方位 2 眼ステレオによる移動物体の位置推定結果

Fig. 16 Estimated object trajectories in a room.

れ、全方位視覚に焦点を絞った国内学会研究会のテーマセッションや国際会議も開かれるようになってきた。従来はロボット視覚への応用が中心であったが、最近では、最終的に人間へのインタラクティブな画像・映像提示を目的としたメディア応用の研究が増える傾向にある。

本論文では、単一視点制約を満たす全方位画像からの任意視線画像の実時間生成による仮想カメラのパン・チルト・ロール操作を利用した、筆者らの研究室でこれまでに行ってきた視覚情報メディアへの全方位ビデオカメラ HyperOmni Vision の応用事例を総合論文としてまとめた。本論文では、テレプレゼンス、代理身体としての移動ロボットの遠隔操縦、およびビデオサーベイランスに話題を限定したが、関連プロジェクトとして、放送型テレプレゼンスシステムの実用化を目指したインタラクティブテレビの実験等も行われている²¹⁾。

HyperOmni Vision に代表される 1 台のカメラを用いて全周の映像を取得する全方位画像センサは、複数カメラを用いる方法 (たとえば、文献 9), 15)) に比べて、機器の構成が簡単で製作しやすい、取得画像からの提示画像の生成が容易である、といった優れた特長を持っている反面、原理的に提示画像の解像度が低いという問題がある。最終的に人間への画像・映像提示を目的とする視覚情報メディア応用では、臨場感を高めるために高解像度化への要求が強い。最近では、ビデオカメラ部にハイビジョンカメラを用いたもの³¹⁾や全方位画像センサを回転あるいは移動させながら取得した複数の画像を用いた超解像度化の試み^{32), 33)}等が

あるが、今後の仮想現実・複合現実応用が期待される大型の没入型投影ディスプレイへの映像提示では、まだ解像度が足りないというのが現状である。全方位画像センサに関しては、さらなる高解像度化が望まれる。

謝辞 本論文で紹介した内容は、ここ数年の修了生が中心となって実施したものである。特に、尾上良雄（現、本田技研工業（株））、三木隆太郎（現、シャープ（株））、山口晃一郎（現、エヌ・ティ・ティ移動通信網（株））、米田美里（現、三洋電機（株））の各氏に感謝する。また、本研究の実施にあたっては、文部科学省・科研費補助金、旧通産省・RWCプログラム、通信・放送機構等の援助を受けた。

参 考 文 献

- 1) Rees, D.W.: Panoramic television viewing system, US Patent No.3505465 (1970).
- 2) Yagi, Y.: Omnidirectional sensing and its applications, *IEICE Trans. Information & Systems*, Vol.E82-D, No.3, pp.568-579 (1999).
- 3) 八木康史: 全方位ビジョンの研究動向, 情報処理学会研究報告, CVIM125-20 (2001).
- 4) 山澤一誠: ミラーを用いた全方位カメラの原理と特徴, 情報処理学会研究報告, CVIM125-21 (2001).
- 5) 横矢直和: 全方位ビデオカメラを用いた広域屋外環境の仮想化, 情報処理学会研究報告, CVIM119-15 (1999).
- 6) 横矢直和: 3D パノラマ画像合成技術, 映像情報メディア学会誌, Vol.54, No.3, pp.338-342 (2000).
- 7) Ohta, Y. and Tamura, H. (Eds.): *Mixed Reality—Merging Real and Virtual Worlds*, Ohmsha & Springer-Verlag, Tokyo (1999).
- 8) Shimamura, J., Yokoya, N., Takemura, H. and Yamazawa, K.: Construction of an immersive mixed environment using an omnidirectional stereo image sensor, *Proc. IEEE Workshop on Omnidirectional Vision*, pp.62-69 (2000).
- 9) 島村 潤, 山澤一誠, 竹村治雄, 横矢直和: 全周パノラマステレオ画像と CG モデルの合成による複合現実環境の構築, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol.42, No.SIG6(CVIM2), pp.44-53 (2001).
- 10) Yamaguchi, K., Takemura, H., Yamazawa, K. and Yokoya, N.: Real-time generation and presentation of view-dependent binocular stereo images using a sequence of omnidirectional images, *Proc. 15th IAPR Int. Conf. on Pattern Recognition*, Vol.4, pp.589-593 (2000).
- 11) Takahashi, T., Kawasaki, H., Ikeuchi, K. and Sakauchi, M.: Arbitrary view position and direction rendering for large-scale scenes, *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Vol.2, pp.296-303 (2000).
- 12) Moezzi, S. (Ed.): Special Issue on Immersive Telepresence, *IEEE MultiMedia*, Vol.4, No.1, pp.17-56 (1997).
- 13) 山澤一誠, 八木康史, 谷内田正彦: 移動ロボットナビゲーションのための全方位視覚系 Hyper-Omni Vision の提案, 電子情報通信学会論文誌 (D-II), Vol.J79-D-II, No.5, pp.698-707 (1996).
- 14) Baker, S. and Nayar, S.K.: A theory of catadioptric image formation, *Proc. 6th Int. Conf. on Computer Vision*, pp.35-42 (1998).
- 15) Kawanishi, T., Yamazawa, K., Iwasa, H., Takemura, H. and Yokoya, N.: Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors, *Proc. 14th IAPR Int. Conf. on Pattern Recognition*, Vol.I, pp.485-489 (1998).
- 16) Nayar, S.K.: Omnidirectional video camera, *Proc. DARPA Image Understanding Workshop*, Vol.1, pp.235-241 (1997).
- 17) Wolberg, G.: *Digital Image Warping*, IEEE Computer Society Press, Los Alamitos, CA (1990).
- 18) 山澤一誠, 尾上良雄, 横矢直和, 竹村治雄: 全方位画像からの視線追従型実時間画像生成によるテレプレゼンス, 電子情報通信学会論文誌 (D-II), Vol.J81-D-II, No.5, pp.880-887 (1998).
- 19) Manabe, Y., Sato, K., Yamazawa, K., Yokoya, N. and Chihara, K.: Reproduction of motion for immersive mixed environments, *Proc. 1st Int. Conf. on Image and Graphics*, pp.673-676 (2000).
- 20) Boulton, T.E.: Remote reality via omnidirectional imaging, *ACM SIGGRAPH'98 Conference Abstracts and Applications*, p.253 (1998).
- 21) 横矢直和: 全方位ビデオカメラの視覚情報メディアへの応用, 情報処理学会研究報告, CVIM125-24 (2001).
- 22) Onoe, Y., Yamazawa, K., Takemura, H. and Yokoya, N.: 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).
- 23) Adelson, E.H. and Bergen, J.: The plenoptic function and the elements of early vision, *Computational Models of Visual Processing*, Landy, M. and Movshon, J.A. (Eds.), pp.3-20, MIT Press, Cambridge, MA (1991).
- 24) Levoy, M. and Hanrahan, P.: Light field ren-

- dering, *Proc. ACM SIGGRAPH'96*, pp.31-42 (1996).
- 25) Gortler, S.J., Grzeszczuk, R., Szeliski, R. and Cohen, M.F.: The lumigraph, *Proc. ACM SIGGRAPH'96*, pp.43-54 (1996).
- 26) 山口晃一郎, 山澤一誠, 竹村治雄, 横矢直和: 全方位動画画像を用いた両眼ステレオ画像の実時間生成によるテレプレゼンス, 電子情報通信学会論文誌 (D-II), Vol.J84-D-II, No.9, pp.2048-2057 (2001).
- 27) 米田美里, 山澤一誠, 竹村治雄, 横矢直和: 全方位動画画像からの両眼ステレオ画像生成による移動ロボットの遠隔操縦, 電子情報通信学会技術研究報告, PRMU2000-107 (2000).
- 28) Nayar, S.K. and Boulton, T.E.: Omnidirectional VSAM system: PI report, *Proc. DARPA Image Understanding Workshop*, Vol.1, pp.55-61 (1997).
- 29) Miki, R., Yokoya, N., Yamazawa, K. and Takemura, H.: A real-time video surveillance and monitoring system using multiple omnidirectional video cameras, *Proc. 4th Asian Conf. on Computer Vision*, Vol.1, pp.528-534 (2000).
- 30) 寺沢征彦, 山澤一誠, 竹村治雄, 横矢直和: 複数の全方位画像センサを用いた遠隔監視システムにおける複数移動物体の存在領域推定, 電子情報通信学会技術研究報告, PRMU2000-195 (2001).
- 31) 山澤一誠, 竹村治雄, 横矢直和: 全方位 HD カメラを用いたテレプレゼンスシステム, 電子情報通信学会技術研究報告, PRMU2001-68 (2001).
- 32) 長原 一, 八木康史, 谷内田正彦: 多重焦点全方位画像列を用いた高精細化, 電子情報通信学会論文誌 (D-II), Vol.J84-D-II, No.8, pp.1882-1890 (2001).
- 33) 川崎 洋, 池内克史, 坂内正夫: 時空間画像解析を用いた全方位カメラ映像の超解像度化, 電子情報通信学会論文誌 (D-II), Vol.J84-D-II, No.8, pp.1891-1902 (2001).

(平成 13 年 7 月 30 日受付)

(平成 13 年 9 月 12 日採録)

(担当編集委員 八木 康史)



横矢 直和 (正会員)

1974 年大阪大学基礎工学部情報工学科卒業。1979 年同大学院基礎工学研究科博士後期課程修了。同年通産省電子技術総合研究所 (現, 産業技術総合研究所) 入所。1986~1987 年カナダ・マギル大学知能機械研究センター客員教授。1993 年奈良先端科学技術大学院大学情報科学センター教授。現在, 同大学情報科学研究科教授, 情報科学センター長 (併任)。画像処理, コンピュータビジョン, 複合現実等の研究に従事。平成元年度情報処理学会論文賞受賞。工学博士。IEEE, 電子情報通信学会, 映像情報メディア学会, 画像電子学会, 人工知能学会, 日本認知科学会, 日本バーチャルリアリティ学会各会員。



山澤 一誠

1992 年大阪大学基礎工学部情報工学科卒業。1994 年同大学院基礎工学研究科博士前期課程修了。1996 年同大学院博士後期課程中退。同年奈良先端科学技術大学院大学情報科学研究科助手。ロボットビジョン, 複合現実の研究に従事。平成 8 年度電子情報通信学会論文賞受賞。博士 (工学)。電子情報通信学会, 日本ロボット学会, 日本バーチャルリアリティ学会各会員。



竹村 治雄 (正会員)

1982 年大阪大学基礎工学部情報工学科卒業。1987 年同大学院基礎工学研究科博士後期課程単位取得退学。同年国際電気通信基礎技術研究所 (ATR) 入社, 通信システム研究所。1994 年奈良先端科学技術大学院大学情報科学研究科助教授。1998~1999 年カナダ・トロント大学客員助教授。2001 年大阪大学サイバーメディアセンター教授。3 次元ユーザインタフェース, CSCW, 仮想現実等の研究に従事。工学博士。IEEE, ACM, HFES, 電子情報通信学会, 映像情報メディア学会, ヒューマンインタフェース学会, 日本バーチャルリアリティ学会各会員。