

NAIST-IS-MT9851118

修士論文

全方位動画像を用いた両眼ステレオ画像の
実時間生成システム

山口 晃一郎

2000年2月14日

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

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学)授与の要件として提出した修士論文である。

山口 晃一郎

審査委員： 横矢 直和 教授
小笠原 司 教授
竹村 治雄 助教授

全方位動画像を用いた両眼ステレオ画像の 実時間生成システム*

山口 晃一郎

内容梗概

遠隔地の情景を画像として提示するテレプレゼンス技術では臨場感豊かに情景を提示することが重要である。その条件として見回しと立体視が可能であることが挙げられる。方向を制御できる可動ステレオカメラを用いることにより遠隔地の情景を広い視野で立体視することは可能であるが、利用者が視線方向を指示してからカメラを回転させるため見回しに関して時間遅延が生じる。本研究では、移動経路と速度が既知である全方位画像センサを用いて撮影した全方位動画像から、任意方向の擬似的な両眼ステレオ画像を光線の情報を用いることにより実時間で生成する手法を提案する。この手法により見回し時の時間遅延なしに、利用者の視線方向の両眼ステレオ画像を提示することが可能となる。

提案手法の有効性を検討するために、提案手法を用いてテレプレゼンスシステムを構築した。システムは一定時間の最新の全方位動画像を計算機に蓄積し、蓄積された全方位動画像から利用者の視線方向に応じた両眼ステレオ画像を生成し、提示する。その結果、両眼ステレオ画像は実時間で生成され、利用者は遠隔地の情景を任意の視線方向で立体視することが可能であることが確認できた。

キーワード

両眼ステレオ画像生成, 全方位画像センサ, HyperOmni Vision, テレプレゼンス, イメージベースレンダリング, plenoptic function

*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 修士論文, NAIST-IS-MT9851118, 2000年2月14日.

Real-time Binocular Stereo Image Generation System Using a Sequence of Omnidirectional Images*

Koichiro Yamaguchi

Abstract

This paper presents a new method to generate and present arbitrarily directional binocular stereo images from a sequence of omnidirectional images in real time. A sequence of omnidirectional images is taken by moving an omnidirectional image sensor in a static real environment. The motion of the omnidirectional image sensor is constrained to a plane. We assume that the sensor's route and speed are known. In the proposed method, a fixed length of the sequence is buffered in a computer to generate arbitrarily directional binocular stereo images using a subset of plenoptic function. Using the method, a user can look around a scene in the distance with rich 3-D sensation without significant time delay. This paper first describes a principle of generating binocular stereo images from a sequence of omnidirectional images. We then describe a telepresence system that generates and presents a pair of view-dependent stereo images in real time, as well as experimental results.

Keywords:

binocular stereo image generation, omnidirectional image sensor, HyperOmni Vision, telepresence, image-based rendering, plenoptic function

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9851118, February 14, 2000.

目次

1. はじめに	1
2. テレプレゼンスの関連研究と本研究の概要	3
2.1 遠隔地の情景提示の従来手法	3
2.1.1 視点位置における画像を獲得する手法	3
2.1.2 任意視点の画像を生成する手法	5
2.2 本研究の概要	9
3. 全方位画像センサ HyperOmni Vision	10
3.1 HyperOmni Vision の特徴	10
3.2 全方位画像と透視投影画像	11
4. 全方位画像系列からの任意視点画像の生成	13
4.1 任意視点画像生成の原理	13
4.2 提案手法の問題点	16
4.2.1 必要な画像枚数の増加	16
4.2.2 垂直方向に生じる歪み	17
4.3 任意視点画像生成の実験	19
4.3.1 実験の概要	21
4.3.2 実験結果	23
5. 見回しと立体視が可能なテレプレゼンスシステム	28
5.1 1入力型実時間テレプレゼンスシステム	28
5.1.1 両眼ステレオ画像生成	29
5.1.2 システム構成	30
5.1.3 実験	32
5.1.4 考察	36
5.2 2入力型蓄積テレプレゼンスシステム	37
5.2.1 両眼ステレオ画像生成	37

5.2.2	システム構成	41
5.2.3	実験	43
5.2.4	考察	47
6.	おわりに	50
	謝辞	51
	参考文献	52

目 次

1	光線の記述	6
2	2点から出る光線の記述	7
3	Concentric Mosaics における光線記述	8
4	パノラマ画像の生成方法	8
5	HyperOmni Vision ver.2A	11
6	全方位画像と透視投影画像	12
7	任意視点画像生成の原理	14
8	垂直方向の疑似生成手法の図解	15
9	経路上に離散的に存在する全方位画像からの任意視点画像の生成	16
10	視線方向と生成に用いる全方位画像の範囲	17
11	縦方向の歪みの生じるときの生成	18
12	物体の距離と縦方向の誤差の関係 (d を変化)	20
13	物体の距離と縦方向の誤差の関係 (視点位置を変化)	20
14	直交する2経路それぞれからの生成に必要な画像の範囲	21
15	任意視点画像の生成位置	22
16	生成された任意視点画像 (1)	25
17	生成された任意視点画像 (2) ~ (7)	26
18	生成された任意視点画像 (8)	27
19	1入力型システムにおける両眼ステレオ画像生成の図解	29
20	1入力型システム構成	31
21	入力全方位動画像	33
22	1入力型システムにおいて生成された両眼ステレオ画像	34
23	生成ステレオ画像に生じる縦視差	35
24	視線方向と生成に必要な画像の範囲の関係 (1入力型システム)	36
25	2入力型システムにおける両眼ステレオ画像生成の図解	38
26	視線方向を変化させたときの両眼ステレオ画像生成	39
27	進行方向に近い光線が必要な場合の生成	39
28	つなぎ目にずれの出る場合の生成	40

29	2 入力型システム構成	42
30	2 経路の入力全方位動画像	44
31	2 入力型システムにおいて生成された両眼ステレオ画像	45
32	生成ステレオ画像に生じる縦方向のずれ	46
33	視線方向と生成に必要な画像の範囲の関係 (2 入力型システム)	48

表 目 次

1	HyperOmni Vision ver.2A パラメータ値	12
2	テレプレゼンスシステムにおける使用機器	30

1. はじめに

遠隔会議や監視システム，移動ロボットの遠隔ナビゲーションなどにおいては，遠隔地の画像を伝送することにより，音声や文字を伝送する場合に比べ，利用者は容易に遠隔地の様子を理解することができる．そのため，こうしたシステムでは遠隔地の画像の獲得に対する要求が高まってきている．ビデオカメラなどの視覚センサを用いて取得した遠隔地の画像を利用者に提示することにより，利用者にその場にいるのと同様の感覚を与える仮想現実技術はテレプレゼンスと呼ばれ，それに関する様々な研究が行われている [1]．テレプレゼンスにおいては，遠隔地の情景を臨場感豊かに提示することが重要となるため，以下の条件を満たすことが望まれる．

(1) 視線変更による自由な見回しが行える：

利用者の視線に追従した画像が時間遅れなく提示され，利用者は遠隔地の情景を見回すことができる．

(2) 両眼立体視が可能である：

遠隔地の情景の両眼ステレオ画像が提示され，利用者は両眼立体視により奥行き感覚を持って遠隔地を観測できる．

遠隔地に設置されたビデオカメラから映像を伝送することにより利用者は遠隔地の様子を知ることができるが，観測できる範囲がビデオカメラのレンズの画角により制限される．また，1つのカメラから撮影された単眼画像では立体視が行えず，奥行き感がない．そこで，カメラを回転台の上に置き，カメラの方向を制御することで観測範囲を広げる方法がとられる [2]．この方法でステレオカメラを用いると，視線方向の両眼ステレオ画像を獲得できるため，利用者は立体視を行うことが可能となるが，見回しに関してはカメラを回転させることによる機械的な時間遅延が発生する．遅延のない見回しを行う方法として，周囲 360 度の情報を含んだ全方位画像を用いる方法が提案されている [3]．全方位の画像から視線方向に応じた画像を提示することにより，利用者は遅延なく見回しを行うことができる．しかし，1つの視点からの画像しか提示できないため立体視を行うこ

とはできない．以上のように，従来の情景提示手法において見回しと立体視を同時に実現することは難しい．

本研究では，撮影位置が既知である全方位画像センサを用いて取得した全方位動画像から，撮影位置付近の任意視点画像を擬似的に生成する手法を提案し，その手法を用いて静止環境で取得された全方位動画像から任意方向の両眼ステレオ画像を実時間で生成・提示するテレプレゼンスシステムを構築する．このシステムを用いると，利用者は遠隔地の情景の時間遅延の少ない見回しと立体視を行うことが可能であり，臨場感豊かに観測を行うことができる．

以下，2章ではテレプレゼンスの関連研究と本研究の概要について述べる．3章では本研究で全方位動画像の取得に用いた全方位画像センサ HyperOmni Vision について説明し，4章では全方位動画像から任意視点画像を擬似的に生成する手法を提案し，その実験について述べる．5章では提案手法を用いて構築した見回しと立体視が可能なテレプレゼンスシステムについて説明し，実験結果と考察を述べる．最後に6章を本論文のまとめとする．

2. テレプレゼンスの関連研究と本研究の概要

テレプレゼンスにおいて臨場感豊かに遠隔地の情景を提示することは重要な課題である。本章では、遠隔地の情景の提示を行う従来手法とその問題点について説明し、本研究の概要を述べる。

2.1 遠隔地の情景提示の従来手法

遠隔地の情景提示を行う従来手法には、カメラの視点位置における画像を獲得して提示を行う手法と、蓄積した画像群から任意の視点位置の画像を生成して提示を行う手法がある。以下、それぞれの従来手法について説明する。

2.1.1 視点位置における画像を獲得する手法

カメラの制御による視点画像の獲得

遠隔地の両眼ステレオ画像を獲得する一般的な方法としては、視線方向が制御できる可動ステレオカメラを使用し、利用者の視線変化に応じてカメラの方向を変更する方法が考えられる。この方法では、ステレオカメラの方向を変えることで利用者の視線に追従した遠隔地の両眼ステレオ画像を取得できるため、利用者に広い視野と奥行き感を与えることが可能である。しかし、視線方向が変化してからカメラを回転させるために、カメラの制御により生じる機械的な遅延と、利用者から遠隔地のカメラまでの通信による遅延が起こり、視線変化から提示される画像の更新までに時間遅延が生じる。

全方位画像からの視点画像の獲得

観測視野を広げるため、周囲 360 度の情報を含んだ全方位画像を取得するさまざまな方法が提案されている [4]。全方位画像の取得方法には、複数の異なる視線方向の画像から全周囲の画像を生成する方法と一度に周囲 360 度の画像を取得する方法がある。

- 複数の画像から生成する方法

QuickTime VR [5] では、カメラを回転させて得られた画像からパノラマ画像を生成しておき、利用者の視線方向に応じた透視投影画像を生成する。しかし、この方法では取得した複数枚の画像をつなぎ合わせてパノラマ画像を生成する必要があるため、全方位動画像の取得は困難である。また、実写画像を撮影した視点位置からしか画像の提示が行えないため立体視をすることはできない。

Virtual Dome [6] では、カメラを常時回転させて全周囲の画像を生成し、その画像から視線方向の画像を切り出して提示する。この方法ではカメラを常時回転させることにより全方位動画像を取得することができるが、カメラの回転に時間を要し、更新レートが低い。また、鏡をカメラの前で回転させる方法 [7] も提案されている。この方法を用いると、カメラ自体を回転させる方法より更新レートを上げることができるが、回転速度を上げるとぶれが生じるために更新レートには限界がある。また、これらの方法は視点位置は固定であるため立体視を行うことはできない。

- 一度に取得する方法

これまでに、観測位置の周囲 360 度の画像を同時に取得できる様々な全方位画像センサが提案されている。全方位画像センサを用いると、一度に全周囲の画像が得られるためにカメラを動かす時間が必要なくなる。全方位画像センサには魚眼レンズを用いる手法 [8]、球面鏡 [9]、円錐鏡 [10]、双曲面鏡 [11]、放物面鏡 [12] などの鏡を用いて全方位画像を取得する手法がある。また、6 角錐鏡と 6 台のカメラを上下対称に配置して高分解能な全方位画像を獲得する全方位ステレオ画像センサも提案されている [13]。これらの手法では全方位画像を一度に撮影できるため、容易に全方位動画像を取得することができる。

全方位画像センサを用いた遠隔地の画像提示手法としては、放物面鏡を用いた全方位画像センサにより任意方向の遠隔地の画像を提示する方法が提案されている [14]。しかし、得られた全方位画像から透視投影画像への正確な変換を行うために生成に時間を要し、利用者の視線方向に追従した

画像を実時間で生成することはできない．一方，双曲面鏡を用いた全方位画像センサから得られた全方位画像を伝送し，実時間テレプレゼンスを行う方法が提案されている [3]．この方法では，生成する疑似透視投影画像を人が歪みを感じない程度にすることで全方位画像からの変換の計算量を減らし，利用者の視線方向に応じた画像を実時間で生成することを可能とした．しかし，これらの手法では全方位画像センサの設置場所を視点とした画像しか生成できないため，立体視を行うことはできない．

2.1.2 任意視点の画像を生成する手法

蓄積した画像群から任意視点位置の画像を生成する手法として，蓄積画像をワーピングやモーフィングにより変換し，画像を生成する手法が提案されている [15, 16, 17, 18]．しかし，これらの手法では蓄積された画像間で各点の対応をとる必要があるため，高速に任意視点の画像を生成することは困難である．

画像間で対応点を求めることなく視点移動を可能とするためには，空間を伝搬するすべての光線の記述が必要となる．3次元空間内を伝搬する光線の記述は Plenoptic Function [19] として提案されている．空間中のすべての光線の情報を記述することで，物体の幾何形状モデルを持たずに新しい視点からの画像を生成することができる．plenoptic とは，ラテン語で「全て」を意味する “plenus” と「光」を意味する “optic” を組み合わせた造語である．以下では，これまでに提案された Plenoptic Function に関する研究を概観する．

- Plenoptic Function [19]

最初に提案された光線データの記述であり $f(x, y, z, \theta, \phi, \lambda, t)$ で表現される7次元の関数で，時間 t に空間中の位置 (x, y, z) から (θ, ϕ) 方向に伝搬する波長 λ の光線の輝度を定義する．

- Plenoptic Modeling [20]

1本の光線は同じ色を持ち，時間により変化しないとすることで λ と t を除いた5次元の Plenoptic Function として提案された．実環境への適用では

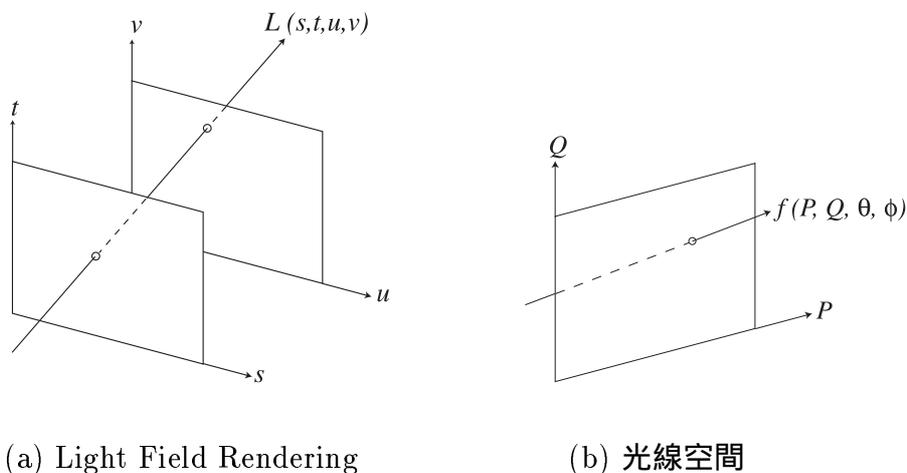


図 1 光線の記述

入力画像が限られ，すべての光線を取得することは困難であるため，画像間の対応付けを行い，光線の得られない部分を補間しなければならない．

- Light Field Rendering [21] , Lumigraph [22] , 光線空間 [23]

光が空間内を伝搬することにより伴う干渉や減衰などの変化を無視することで関数を 5 次元から 4 次元に射影した．図 1 に示すように，Light Field Rendering は光線を 2 つの平面を通る点の位置 (s, t) と (u, v) により，光線空間は 1 つの平面を通る点の位置 (P, Q) と方向 (θ, ϕ) により光線を記述するが，原理的には同等のものと考えることができる．また Lumigraph では，基本原理として Light Field Rendering を採用し，任意視点画像を生成するシステムを構築した．ここで，図 2 において矢印で示す点 A から出る光線と点 B から出る光線について考える．光線の方向を (θ, ϕ) とすると，5 次元の記述では，点 A から出る光線は $f(x_a, y_a, z_a, \theta, \phi)$ で，点 B から出る光線は $f(x_b, y_b, z_b, \theta, \phi)$ となり異なるが，Light Field Rendering では光の伝搬に伴う変化を無視するためどちらの光線も $L(s, t, u, v)$ と記述され，点 A からの光線と点 B からの光線が同等となる．したがって，これらの手法では新しい視点の画像に必要な光線と同等の光線を含む画像群を取得すれば，すべての光線の記述を求めることなく，新しい視点の画像を生成することが

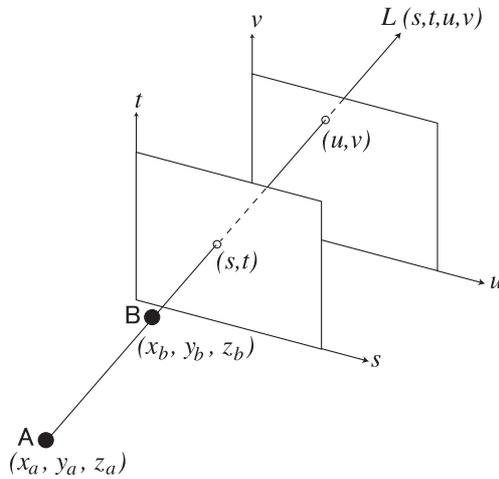


図 2 2点から出る光線の記述

できる．そのため，5次元での光線記述を行う場合よりも画像枚数を減らすことができるが，任意視点画像の生成には多数の画像を必要とするためにデータ量が膨大となる問題がある．

- Concentric Mosaics [24]

この手法では，新しい視点位置を2次元平面上の円の中に制限することで光線の関数を3次元とした．図3に示すように，光線を半径 r ，角度 ρ ，光線が平面となす垂直方向の角度 ϕ で記述する．この記述により平面上の円の中の点から出るすべての光線には，ある半径の円周上のある点から接線方向へ出る同等の光線が存在する．必要な光線を取得するために，図4のように円周に沿ってカメラを移動させて得られた画像から Manifold Mosaic [25] と呼ばれるパノラマ画像を生成する．このパノラマ画像により，円周上の点から接線方向への光線を得ることができる．同様の方法で生成された同心円のパノラマ画像を複数枚用いることで，異なる半径の円周からの光線を取得する．視点位置を2次元平面上の円の中に制限することにより，Light Field Rendering などに比べて必要なデータ量も少ない．この手法を用いれば平面上の円の中に位置する任意視点画像を生成することは可能であるが，

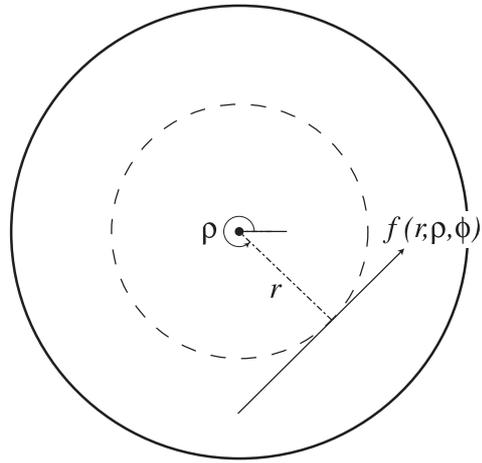


図 3 Concentric Mosaics における光線記述

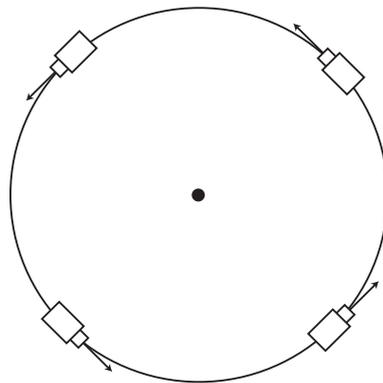


図 4 パノラマ画像の生成方法

カメラを回転させて得られた画像から生成するパノラマ画像が複数枚必要となるため、あらかじめ画像の取得を行っておかなければならない。

以上の光線情報を用いて遠隔地の情景提示を行うことを考えたとき、任意方向の両眼ステレオ画像を奥行き情報を求めることなく生成することが可能となるが、多数の画像が必要となりデータ量が増加することや、あらかじめ画像群を取得しておく必要があるため、実時間テレプレゼンスが困難であるという問題がある。

2.2 本研究の概要

本研究では、静止環境下で見回しと立体視が可能なテレプレゼンスを行うことを目的とし、全方位動画像を用いて利用者の視線方向の両眼ステレオ画像を実時間で生成し、提示するテレプレゼンスシステムを構築する。全方位動画像を伝送し、光線情報を用いて利用者の視線方向の両眼ステレオ画像を生成、提示するため、利用者は少ない時間遅延で見回しと立体視を行うことができる。任意方向の両眼ステレオ画像を生成するため、撮影位置が既知である全方位動画像から光線情報を用いて任意視点画像を生成する手法を提案する。提案手法では、周囲 360 度の情報を含む全方位画像を用いること、Light Field Rendering などと同様に光が伝搬することにより伴う減衰などの変化を無視すること、生成する画像の視点位置を経路付近に制限することにより少ないデータ量で任意視点画像を生成する。また、撮影位置が既知である全方位動画像は全方位画像センサを等速で直線移動させて取得し、センサの移動速度は既知とする。全方位画像センサには、取得された全方位画像から幾何学変換により平面透視投影画像に変換することのできる HyperOmni Vision [11] を用いる。

テレプレゼンスシステムは、1 つの全方位動画像を入力とする 1 入力型と、2 つの全方位動画像を同時に入力する 2 入力型を作成した。1 入力型システムでは、利用者の視線方向の範囲が制限されるが、実時間でテレプレゼンスを行うことが可能である。一方、2 入力型システムでは任意の方向の両眼ステレオ画像を提示することができるが、眼間距離だけ離れ平行な 2 経路の全方位動画像が必要となる。そのため、全方位画像センサが眼間距離より大きい場合には同時に取得できず、あらかじめ全方位動画像を撮影する必要がある。

3. 全方位画像センサ HyperOmni Vision

これまでに、観測位置において周囲 360 度の画像を同時に取得できるさまざまな全方位画像センサが提案されているが、本研究では全方位動画像を取得するために用いる全方位画像センサとして、HyperOmni Vision [11] を採用した。HyperOmni Vision は双曲面鏡を用いた全方位画像センサであり、センサの周囲 360 度の画像を一度に取得できる。また、他の多くの全方位画像センサとは異なり、光学系の特性が透視投影であるため、得られる全方位画像から鏡の焦点から見た通常の平面透視投影画像を生成できる。以下では、HyperOmni Vision の構成ならびに得られる全方位画像の特徴について述べる。

HyperOmni Vision の外観と光学系の構成を図 5 に示す。本センサは、図 5(a) に示すように、鉛直下向きに設置した双曲面鏡に写り込んだ鏡像を、鉛直上向きに設置したビデオカメラで撮影するものである。双曲面には回転 2 葉双曲面を用いており、図 5(b) のように 2 焦点 O_M, O_C を持っている。鏡の焦点とカメラのレンズ中心は、それぞれ O_M, O_C に位置している。ここで図 5(b) に示すように Z 軸を鉛直軸とする 3 次元ワールド座標系 (X, Y, Z) を考えると、双曲面鏡の面、双曲面鏡の焦点 O_M 、ならびにカメラのレンズ中心 O_C は式 (1) で表せる。本研究で使用した HyperOmni Vision ver.2A の仕様を表 1 に示す。

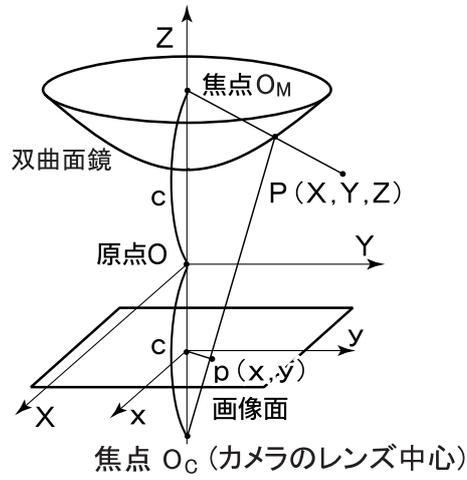
$$\begin{aligned} \text{鏡面} & : \frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 \\ & \quad (Z > 0) \\ \text{鏡の焦点 } O_M & : (0, 0, +c) \\ & \quad (c = \sqrt{a^2 + b^2}) \\ \text{カメラレンズ中心 } O_C & : (0, 0, -c) \\ \text{ただし } a, b, c & : \text{ 双曲面鏡のパラメータ} \end{aligned} \tag{1}$$

3.1 HyperOmni Vision の特徴

内側の焦点 O_M に集まる像は双曲面鏡を介して、外側の焦点 O_C に集まる。従って、 O_C にレンズ中心を置いたカメラで、双曲面鏡の焦点 O_M への全方位の双曲面中心投影像 (の鏡像) を撮影することができる。



(a) 外観



(b) 光学系の構成

図 5 HyperOmni Vision ver.2A

双曲面鏡の焦点 O_M とカメラのレンズ中心 O_C との位置関係ならびに双曲面の特性により，3次元環境中の点 $P(X, Y, Z)$ と全方位画像上の写像点 $p(x, y)$ の間には式 (2) の関係が成立する (図 5(b) 参照) .

$$x = \frac{X f (b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}$$

$$y = \frac{Y f (b^2 - c^2)}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}} \quad (2)$$

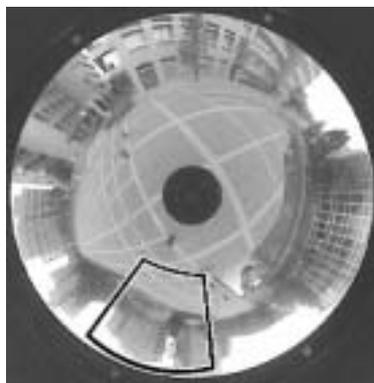
3.2 全方位画像と透視投影画像

HyperOmni Vision により得られる全方位画像の例を図 6(a) に示す . 同図のように，双曲面鏡の焦点 O_M を中心として周囲 360 度の画像が一度に取得できる .

全方位画像は双曲面鏡の焦点 O_M を投影中心とした双曲面上への 1 点中心投影像であるため，その一部を通常の平面透視投影像に変換することができる . 図 6(a) の一部を変換して通常の平面透視投影画像を生成した例を図 6(b) に示す . こ

表 1 HyperOmni Vision ver.2A パラメータ値

a	42.1 mm
b	42.7 mm
c	59.1 mm
鏡の直径	150 mm
カメラの焦点距離 f	5.4–64.8 mm
撮像素子	1/3" IT CCD
センサ視野	水平より上方視野 20 度，下方視野 90 度



(a) 全方位画像



(b) 透視投影画像

図 6 全方位画像と透視投影画像

の透視投影画像は，図 6(a) の全方位画像上の四角形の枠と対応している．この変換は透視投影画像のすべての画素に対応する全方位画像上の対応点の位置を求めることにより行った．図 6(a) の全方位画像では周囲 360 度すべての情報を含むため画像が歪んだものとなるが，図 6(b) の透視投影画像では歪みのない画像となっていることが分かる．

4. 全方位画像系列からの任意視点画像の生成

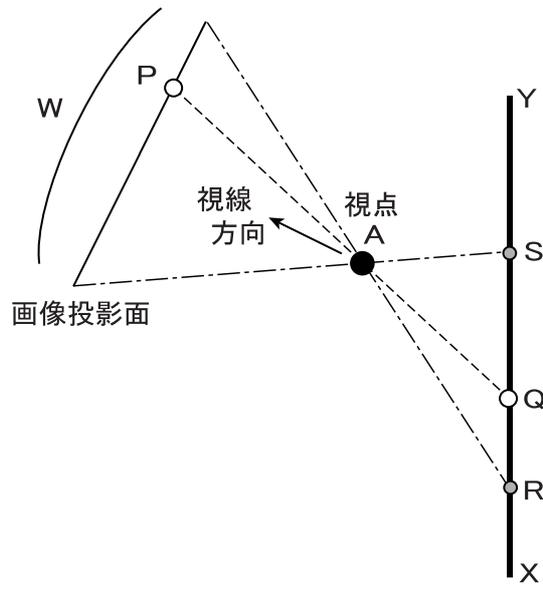
全方位画像センサを水平に直線移動させて取得された全方位動画像を用いて，移動経路を含む水平面に位置する新しい視点の画像を生成する手法を提案する [26, 27]．この提案手法と同原理の手法が，高橋らによって提案されている [28]．以下では，提案手法の原理と問題点について説明し，任意視点画像生成の実験について述べる．

4.1 任意視点画像生成の原理

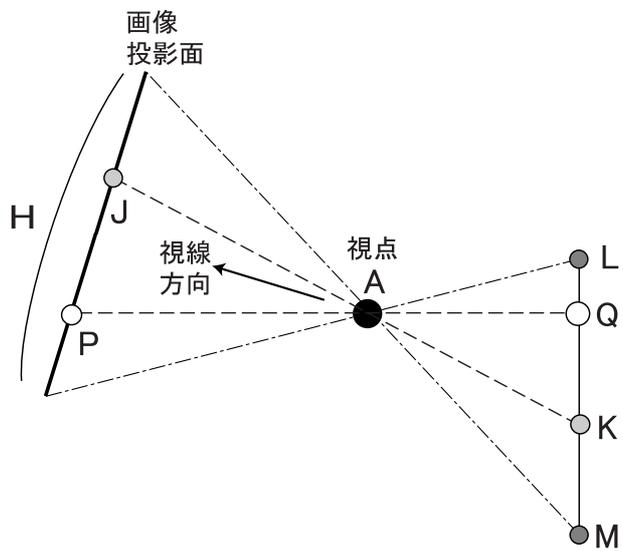
本研究では，全方位動画像は静止環境において全方位画像センサ HyperOmni Vision を等速で水平に直線移動させながら撮影することにより取得し，センサの移動速度は既知であるとする．また，生成する画像の視点は全方位画像センサを移動させた水平面に位置しているとする．

例えば HyperOmni Vision を図 7(a) の地点 X から地点 Y まで直線移動させ全方位動画像を取得し，視線が矢印の方向の視点 A の画像 (解像度 $W \times H$ 画素) を生成することを考える．このとき生成する視点の画像の投影面は図 7 に示す位置となる．図 7(b) は QP を含む垂直な面を横から見た図である．光が伝搬により変化しないとすると，図 7 のような位置関係のとき新しい視点位置から出る光線 AP と，直線 \overline{AP} と直線 \overline{XY} の交点 Q から出る光線 QP を同等とすることができる [21, 22, 23] ので，点 Q で撮影された全方位画像から生成画像の点 P の画素値を求めることができる．また，図 7(b) に示すように光線 AJ と光線 KJ は同等であるので点 K で取得される画像から点 J の画素値を得ることができる．同様に考えると，図 7(b) で示す生成画像の点 P を含む高さ H 画素の 1 列は点 L から点 M の範囲で取得される画像から生成できる．しかし，全方位動画像は全方位画像センサを点 X から Y まで水平に直線移動させて取得したものであり，図 7(b) で示す点 L から点 M の範囲で点 Q 以外の点では全方位画像は取得されていない．そのため提案手法では，生成画像の点 P を含む高さ H 画素の 1 列すべてを点 Q の全方位画像を用いて擬似的に生成する．

画像の点 J を生成するには光線 AJ と同等の光線が必要となるが，光線 AJ と



(a) 上から見た図



(b) 横から見た図

図 7 任意視点画像生成の原理

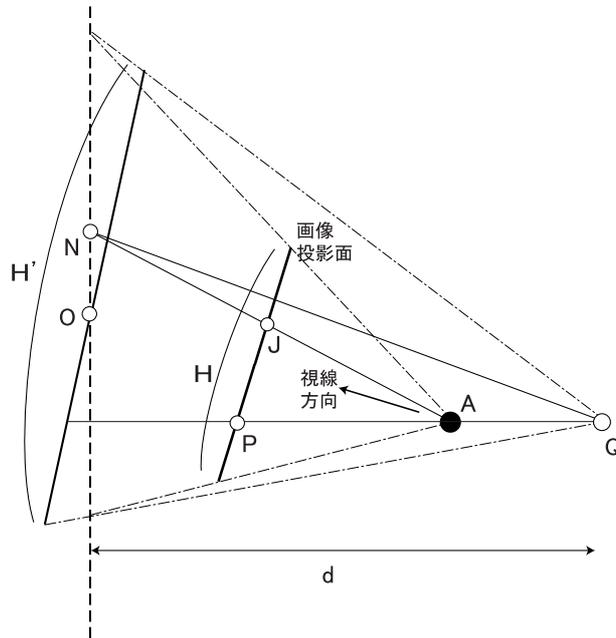


図 8 垂直方向の疑似生成手法の図解

光線 QJ は異なる光線であり点 Q からの光線で光線 AJ と同等の光線はない。そこで提案手法では、図 8 で示すように点 Q からの距離が d である垂直線上で光線 AJ と交わる光線 QN を光線 AJ と同等とすることにより擬似的に画像の生成を行う。同様に、生成画像の点 P を含む高さ H の 1 列の生成に必要なすべての光線を、点 Q からの距離が d である垂直線上でそれぞれの光線と交わる光線として生成を行う。図 8 において、視点位置 A から、点 P を含む高さ H の 1 列の midpoint への光線と同等とする光線は光線 QO となる。このとき、点 Q の全方位画像から点 O へ方向に焦点距離 $|QO|$ として幅 1、高さ H' 画素の透視投影画像を生成し、それに縦方向のスケーリングを行い幅 1、高さ H 画素の画像とすることで擬似的に画像の 1 列を生成する。

同様の方法により、図 7(a) に示す地点 R から地点 S までの全方位画像を用いて点 A に視点位置する画像を生成する。しかし、生成に用いる全方位動画は 1 秒当たり 30 枚の全方位画像から構成されており、図 9 に示す T_1 から T_5 の

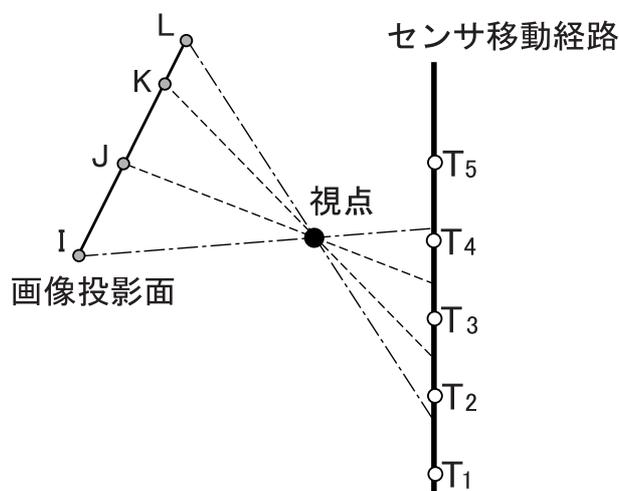


図 9 経路上に離散的に存在する全方位画像からの任意視点画像の生成

ように、移動経路上で全方位画像は離散的に存在しているために必要な地点における全方位画像が存在しない場合がある。その場合には、存在する全方位画像の中で必要な地点に最も近い位置の全方位画像を用いて画像の生成を行うこととする。図 9 の場合、生成画像の点 I から点 J までの列は T_4 、点 J から点 K までの列は T_3 、点 K から点 L までの列は T_2 の全方位画像を用いて生成することになる。

4.2 提案手法の問題点

4.2.1 必要な画像枚数の増加

前節で述べた任意視点画像生成では、直線経路上の複数枚の全方位画像を用いて生成を行う。図 10 の視線方向 1 のように視線方向が経路と直交する方向に近い場合には問題とならないが、図 10 の視線方向 2 のように視線方向が経路と平行に近くなると、生成に用いる全方位画像の範囲、点 T_2 から点 U_2 が、視線方向 1 の場合の点 T_1 から点 U_1 と比べて長くなり、必要な画像枚数が大きく増加する。つまり、視線方向が経路に平行に近づいた場合、任意視点画像の生成に大量の全方位画像を必要とする。

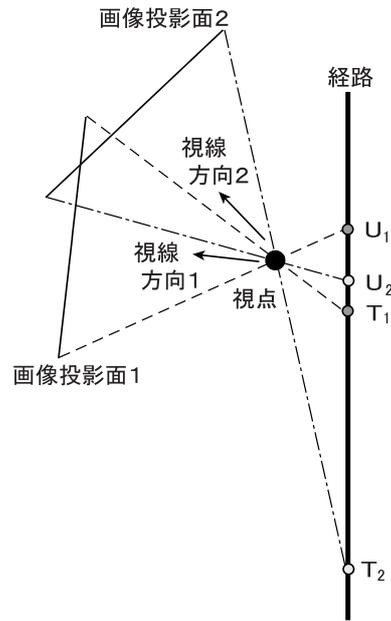


図 10 視線方向と生成に用いる全方位画像の範囲

4.2.2 垂直方向に生じる歪み

前節で述べた方法により任意視点画像を生成した場合，直進する全方位画像を用いているため生成画像の水平方向に関しては複数の全方位画像から画像計算が行われるが，垂直方向に関しては同じ全方位画像から生成されるため擬似的な生成となり，縦方向に歪みが生じる．撮影環境内のすべての物体の形状，位置が既知であればこの歪みを補正することは可能であるが，本研究では撮影環境は未知であるので，撮影位置から物体までの距離を一定 (d) と仮定することで生成を行っている．

この生成方法では，すべての物体までの距離を一定としているために図 11 のような d と異なる位置に物体が存在している場合には縦方向に歪みが生じる．図 11 の場合，生成したい視点の位置から投影された画像には物体の点 M_a から点 N_a までが写ることになるが，全方位画像から撮影位置から物体までの距離を d とし生成した画像には物体の点 M_b から点 N_b までしか写らずに，生成された画像は縦方向に伸ばされたものとなる．図 11 のように，物体に対して全方位画像撮影

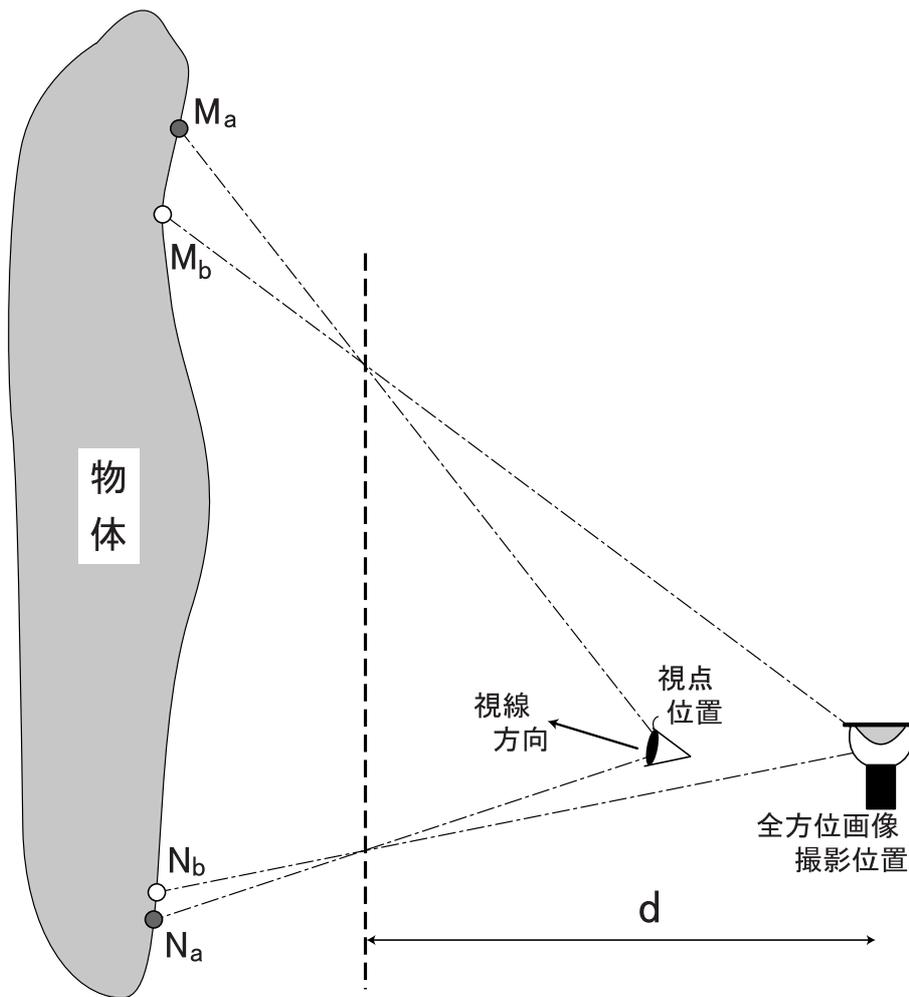


図 11 縦方向の歪みの生じるときの生成

位置が視点よりも遠い位置にあり，実際の物体位置が d よりも遠いときには生成画像は縦方向に伸ばされ，逆に，実際の物体の位置が d よりも近いときには生成画像は縦方向に縮められる．また，全方位画像撮影位置が視点よりも物体に近い場合にはその逆となる．つまり，誤差(歪み)は d と物体までの実際の距離の差と，全方位動画像の経路と視点の位置の距離に依存する．

d を 1m , 2m , 3m , 4m , 5m , 無限遠とした場合に撮影位置から物体までの実際の距離と生成画像に生じる誤差の関係を図 12 に示す．図 12 は，視線方向は全方位動画像の進行方向に対して左に 45 度，垂直方向に 0 度，視点位置は全方位動画像の経路からの距離が 3cm ，生成する画像は水平画角を 50 度，解像度を 640×480 画素とした場合を示している．また縦方向のずれは生成画像中の位置により変化するが，図 12 の縦軸は生成画像中で最大の誤差となるとき画素数で，横軸は実際の物体の距離を示している． d を無限遠とした場合は，新しい視点位置からの垂直画角と同じ垂直画角で全方位画像から生成を行う．図 12 に示すように， d を変えることにより生成画像における縦方向の誤差も変化し， d と実際の距離が離れるにつれ誤差は大きくなる．図 12 の物体の距離が 3m のときに注目すると， d が 2m の場合の方が d が 4m の場合よりも誤差が大きく， d を実際の距離より小さくした場合は大きくした場合よりも誤差が大きくなることが分かる．

縦方向のずれは全方位動画像の経路から視点位置の距離によっても変化する．図 13 は，経路から視点位置までの距離を 3cm , 5cm , 10cm , 15cm , 20cm としたときに生じる縦方向の誤差を示している．図 13 で示すのは，視線方向は全方位動画像の進行方向に対して左に 45 度，垂直方向に 0 度， d は 3m ，生成する画像は水平画角が 50 度で，解像度が 640×480 画素の場合である．図 13 から，視点位置を経路から離していくにしたがって最大誤差も大きくなっていくことが分かる．したがって，提案手法により生成する画像の視点位置は経路付近に制限する必要がある．

4.3 任意視点画像生成の実験

本節では，提案手法を用いて，全方位動画像から任意視点画像を生成する実験とその結果について述べる．

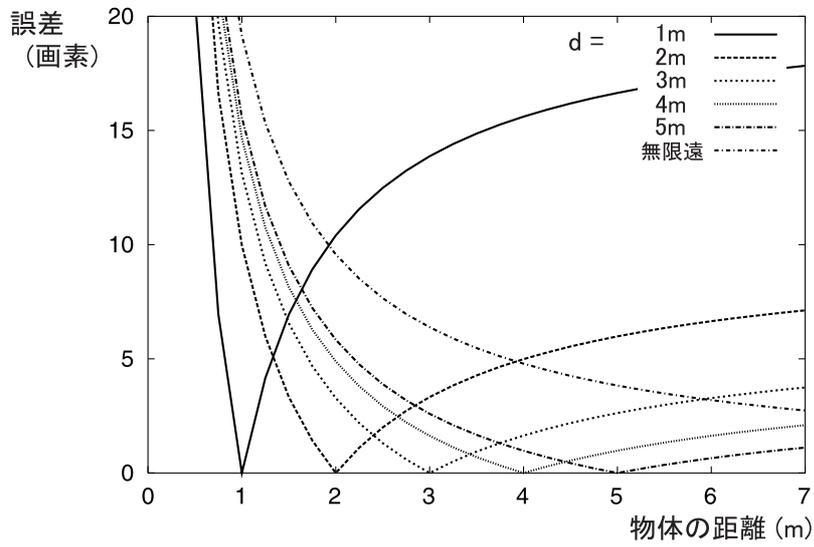


図 12 物体の距離と縦方向の誤差の関係 (d を変化)

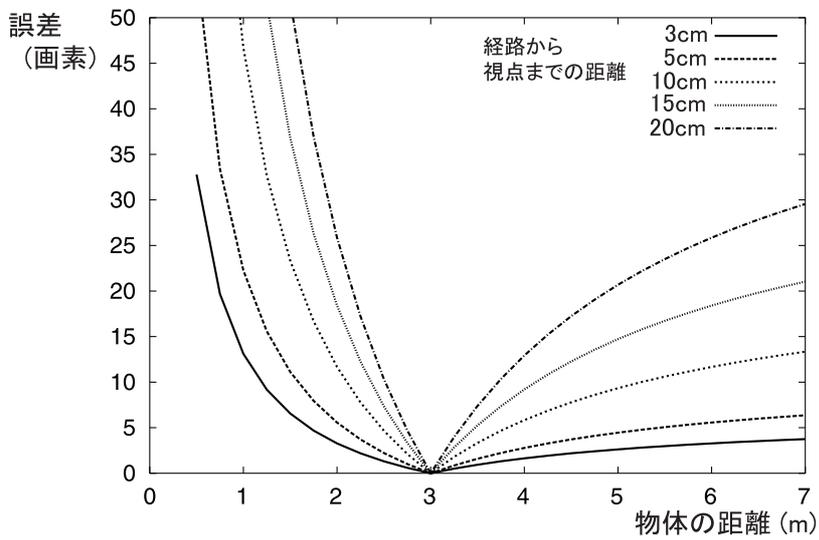


図 13 物体の距離と縦方向の誤差の関係 (視点位置を変化)

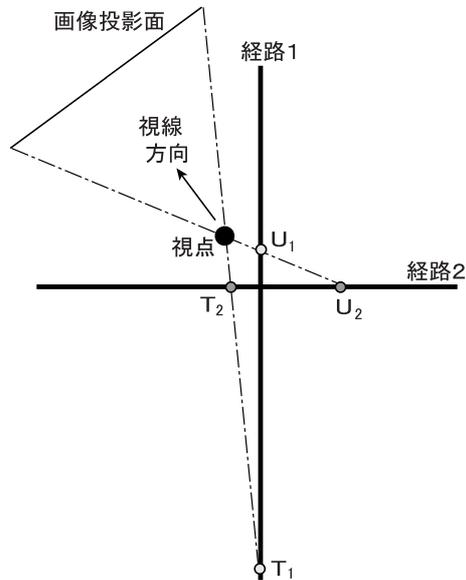


図 14 直交する 2 経路それぞれからの生成に必要な画像の範囲

4.3.1 実験の概要

提案手法では，前述したように視線方向が経路と平行に近づいたとき，生成に必要な画像枚数が大きく増加する．図 14 のように直交する 2 経路，経路 1 と経路 2 の全方位動画像を取得し，経路 1 に平行に近い視線方向の画像を生成するとき，画像生成に経路 1 の全方位動画像を用いると T_1 から U_1 までの長い範囲に存在する全方位画像が必要となるが，経路 1 に直交する経路 2 の全方位動画像を用いると T_2 から U_2 までの短い範囲に存在する全方位画像から生成できる．つまり，図 14 のように直交する 2 経路を取得すれば，大量の全方位画像を用いることなく，すべての視線方向に対して生成を行うことができる．そこで，任意の位置からの任意方向の画像を生成する実験を行うため，全方位動画像は図 15 に示すように直交する 2 つの経路で取得した．全方位動画像は，HyperOmni Vision を載せた移動ロボット Nomad-200 をそれぞれの経路で 17.5cm/s の一定速度で 6 秒間直進させることにより取得した．つまり，それぞれの経路上に 5.8mm 間隔で全方位画像 180 枚が存在している．任意視点画像の生成には SGI Onyx2 IR2 を使用し，全方位動画像をすべてメモリに格納して生成を行った．

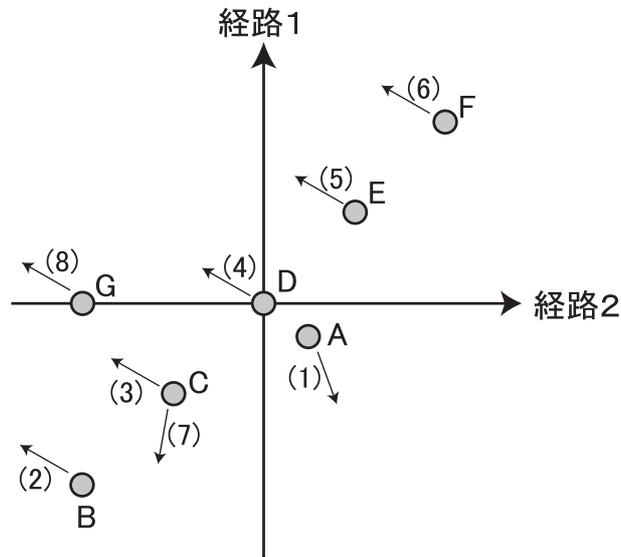


図 15 任意視点画像の生成位置

実験では図 15 に示す点 A から点 G までの 7 地点からの視点において矢印の示す視線方向の画像 (1) から (8) を生成した。画像 (1) の視線方向は経路 1 の進行方向から右に 160 度で視点位置は点 A である。画像 (2) から (6) はすべて視線方向が経路 1 の進行方向から左に 60 度で、視点の位置はそれぞれ点 B から点 F である。画像 (7) は視点位置は点 C で、視線方向は経路 1 の進行方向から左に 170 度である。画像 (8) は視点位置が点 G で、視線方向は経路 1 の進行方向から左に 60 度である。点 D は経路 1 と経路 2 が直行する点であり、点 A は点 D から経路 1 の進行方向の逆方向へ 1.7cm、経路 2 の進行方向へ 1.2cm 移動した点である。点 E は点 D から経路 1 と経路 2 の進行方向にそれぞれ 9.9cm 移動した点で、点 F はさらに点 E から経路 1 と経路 2 の進行方向にそれぞれ 9.9cm 移動した点である。点 C、B は点 E、F と逆方向に移動した点である。点 G は点 D から経路 2 の進行方向と逆方向へ 19.8cm 移動した点である。

4.3.2 実験結果

図 16の下に示す画像が生成された画像 (1) である。画像 (1) は経路 2 の全方位動画像を用いて生成した。図 16に示すように、画像 (1) は経路 2 の全方位動画像中の 4 枚の全方位画像 O_1 から O_4 から生成される。図 16の全方位画像 O_1 から O_4 の四角形の枠は画像 (1) の生成に用いている部分を示している。

図 17に生成した画像 (2) から (7) を示す。画像 (2) から (6) の画像に存在する椅子の角度の違いから、視点の位置が変化している画像が生成されていることが確認できる。画像 (2) から (6) は経路 1 の全方位動画像から生成し、画像 (7) は経路 2 の全方位動画像から生成した。前述したように、本手法では全方位動画像の進行方向に近い視線方向の画像を生成するとき、多くの全方位画像を必要とするために生成が困難となる。そこで実験では、視線方向が経路 1 の進行方向に対して左右に 45 度から 135 度までの画像は経路 1 の全方位動画像を用いて生成し、それ以外の視線方向の画像は経路 2 の全方位動画像を用いて生成することにより任意方向の画像を生成した。また、図 16, 17に示した画像 (4) を除く画像は、撮影位置から物体までの仮定距離 d を 3m として生成し、画像 (4) は経路上に視点があるため、点 D に存在する 1 枚の全方位画像から生成した。

図 18に画像 (8) を異なる方法で生成した 4 枚の画像を示す。図 18(a) は経路 2 の全方位動画像中に存在する点 G の全方位画像 1 枚から生成した画像である。図 18(b), (c), (d) は経路 1 の全方位動画像を用いて生成した画像で、(b), (c), (d) で撮影位置から物体までの仮定距離 d を変えて生成している。(b) では d を 1m, (c) では d を 3m, (d) では d を無限遠として生成した。図 18(a) は実際に点 G で取得された全方位画像から生成しているために歪みのない画像であるが、任意視点画像生成手法により生成された (b), (c), (d) には縦方向に歪みが生じる。画像内に存在する物体は経路 1 の撮影位置から約 2.5m から約 5m までの範囲に存在する。図 18(b), (c), (d) それぞれと (a) を比較すると、 d を 1m として生成した (b) では、画像は大きく縦方向に歪んでおり、縦方向に伸ばされていることが分かる。 d を無限遠として生成した (d) は画像内の物体のうち近くにある椅子に注目すると、画像が縦方向に縮められていることが分かる。 d を 3m とした (c) の画像は (b), (d) の画像に比べ、縦方向の歪みが小さくなっていることが確認で

きる .

図 17, 18で画像 (2), (6) は 46 枚, (3), (5) は 24 枚, (4) は 1 枚, (7) は 17 枚, (8) の (a) は 1 枚, (8) の (b), (c), (d) は 46 枚の全方位画像を用いて生成された . また, 生成に用いる画像がすべてメモリに読み込まれてから, 任意視点画像 1 枚を生成するのに要した時間は 770 ミリ秒であった .

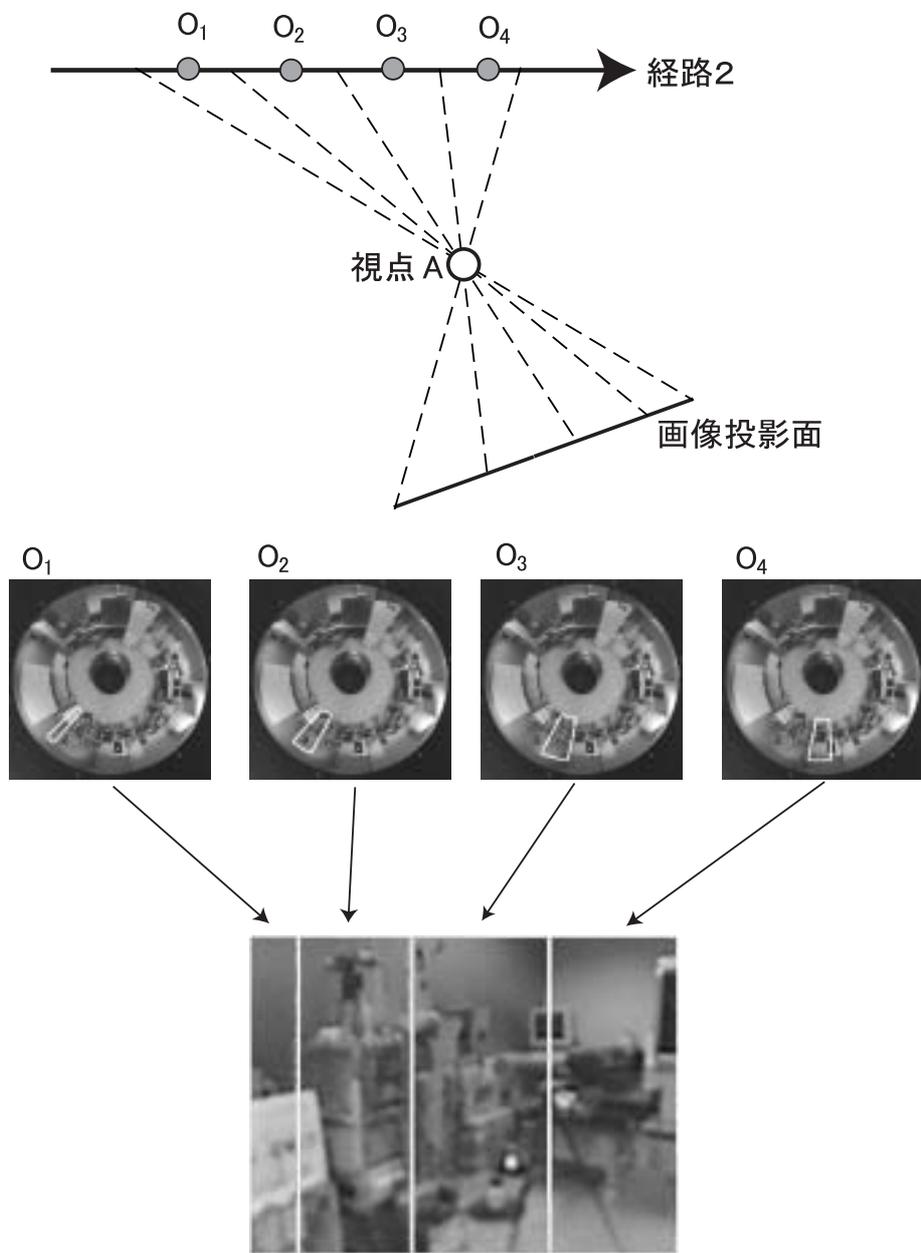


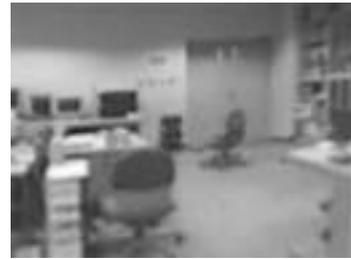
図 16 生成された任意視点画像 (1)



画像 (2)



画像 (3)



画像 (4)



画像 (5)



画像 (6)



画像 (7)

図 17 生成された任意視点画像 (2) ~ (7)



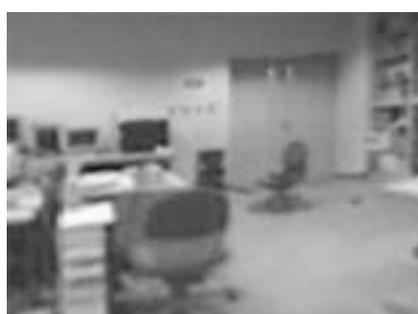
(a) 歪みなし



(b) $d = 1\text{m}$



(c) $d = 3\text{m}$



(d) $d = \infty$

図 18 生成された任意視点画像 (8)

5. 見回しと立体視が可能なテレプレゼンスシステム

本章では、前章で述べた任意視点画像生成手法を用いて、全方位画像センサを一定速度で水平に直進させて取得した全方位動画像から任意方向の両眼ステレオ画像を実時間で生成・提示するテレプレゼンスシステムについて述べる。このシステムでは一定時間の全方位動画像を計算機に保持させ、その全方位動画像から両眼ステレオ画像を生成する。

システムは、1つの全方位動画像から両眼ステレオ画像を生成する1入力型システムと2つの全方位動画像から両眼ステレオ画像を生成する2入力型システムを構築した。1入力型システムは、保持できる画像枚数の制約のために、視線方向により両眼ステレオ画像が生成できない場合があるが、実時間テレプレゼンスを行うことが可能である。一方、2入力型システムでは、2つの全方位動画像を用いることにより全方位の視線方向の両眼ステレオ画像を生成することを可能としている。しかし、眼間距離だけ離れ平行な2経路の全方位動画像が必要となり、眼間距離より大きい全方位画像センサを用いるときには、同時に2つの全方位動画像を取得できないため、全方位動画像をあらかじめ取得する蓄積型テレプレゼンスとなる。以下では、それぞれについてシステムの概要、実験そして考察を述べる。

5.1 1入力型実時間テレプレゼンスシステム

1入力型システムでは、全方位画像センサを直進させることにより取得した1つの全方位動画像を計算機に伝送し、任意方向の両眼ステレオ画像を実時間で生成・提示する。このシステムでは直線経路の全方位動画像1つから両眼ステレオ画像を生成するため、経路に対して平行に近い視線方向の両眼ステレオ画像の生成には多数の全方位画像が必要となり生成が困難となる。しかし、1つの全方位動画像を用いるため、伝送された全方位動画像から両眼ステレオ画像を実時間で生成することにより実時間テレプレゼンスが可能である。

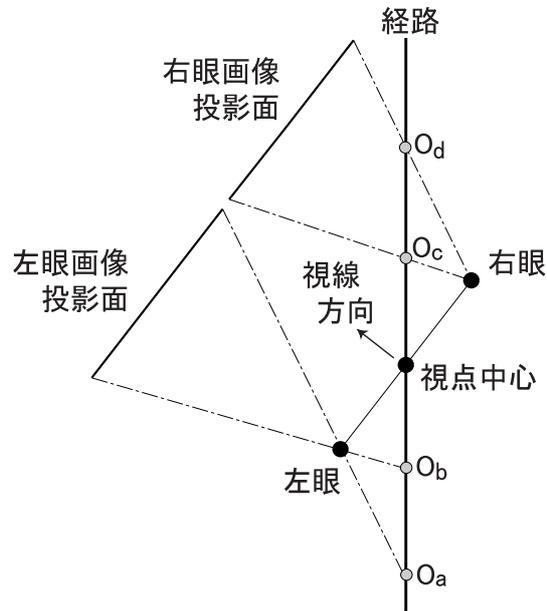


図 19 1 入力型システムにおける両眼ステレオ画像生成の図解

5.1.1 両眼ステレオ画像生成

生成したい両眼ステレオ画像の左眼と右眼が全方位動画像の経路を含む水平面上にあれば前章の任意視点画像生成手法を用いて左眼画像，右眼画像を生成することができる．このシステムでは図 19 のように視点中心が経路上にある両眼ステレオ画像を生成する．図 19 に示す位置の視点中心から視線方向の両眼ステレオ画像を生成するとき，左眼と右眼は図 19 に示す位置となる．任意視点画像生成手法を用いると左眼画像は全方位動画像中の O_a から O_b の範囲，右眼画像は O_c から O_d の範囲の全方位画像から生成できる．

テレプレゼンスシステムにおいては視線方向に追従した両眼ステレオ画像を提示するため，実時間で画像生成を行う必要がある．そこで，両眼ステレオ画像生成において，生成画像全点に対して全方位画像との対応計算を行わず，生成画像に格子を設け，格子点でのみ対応計算を行い，格子点間はテクスチャマッピング機能を利用してハードウェアによる画像変形 [29] を行う手法 [3] を用いて高速に画像の生成を行う．

表 2 テレプレゼンスシステムにおける使用機器

全方位視覚センサ	HyperOmni Vision ver.2A
移動ロボット	Nomad-200
計算機	SGI Onyx2 IR2 (MIPS R10000 195MHz, 16CPU)
磁気トラッカ	POLHEMUS 3SPACE FASTRAK
HMD	OLYMPUS Mediamask

また 4.2.1 節で述べたように，視線方向が経路と平行に近くなると両眼ステレオ画像の生成には視点中心から離れた多数の全方位画像が必要となる．このシステムでは一定時間の全方位動画像を計算機に保持させるため，視点中心から遠く離れた全方位画像が必要となる場合には両眼ステレオ画像が生成できなくなる．また任意視点画像生成手法では生成画像に縦方向の歪みが生じるが，視点中心が全方位動画像の経路上にあるとき左眼画像に生じる縦歪みと右眼画像に生じる縦歪みは逆になり，生成された両眼ステレオ画像には縦方向の視差が生じる．

5.1.2 システム構成

1 入力型システムの構成を図 20 に，使用した機器を表 2 に示す．システムでは，全方位画像センサ HyperOmni Vision を載せた移動ロボットを遠隔地で移動させて全方位動画像を取得する．利用者は HMD を装着し，遠隔地の情景の見回しと立体視を行うことができる．

1 入力型システムにおける処理の流れは以下の通りである．

1. 全方位画像センサ HyperOmni Vision を移動ロボットに載せ，計算機から制御する移動ロボットを一定速度で直進させながら撮影した全方位動画像を計算機に伝送し，計算機は一定時間の全方位動画像を保持する．
2. 利用者の装着する HMD に取り付けた 3D 磁気トラッカで視線 (頭部の向き)

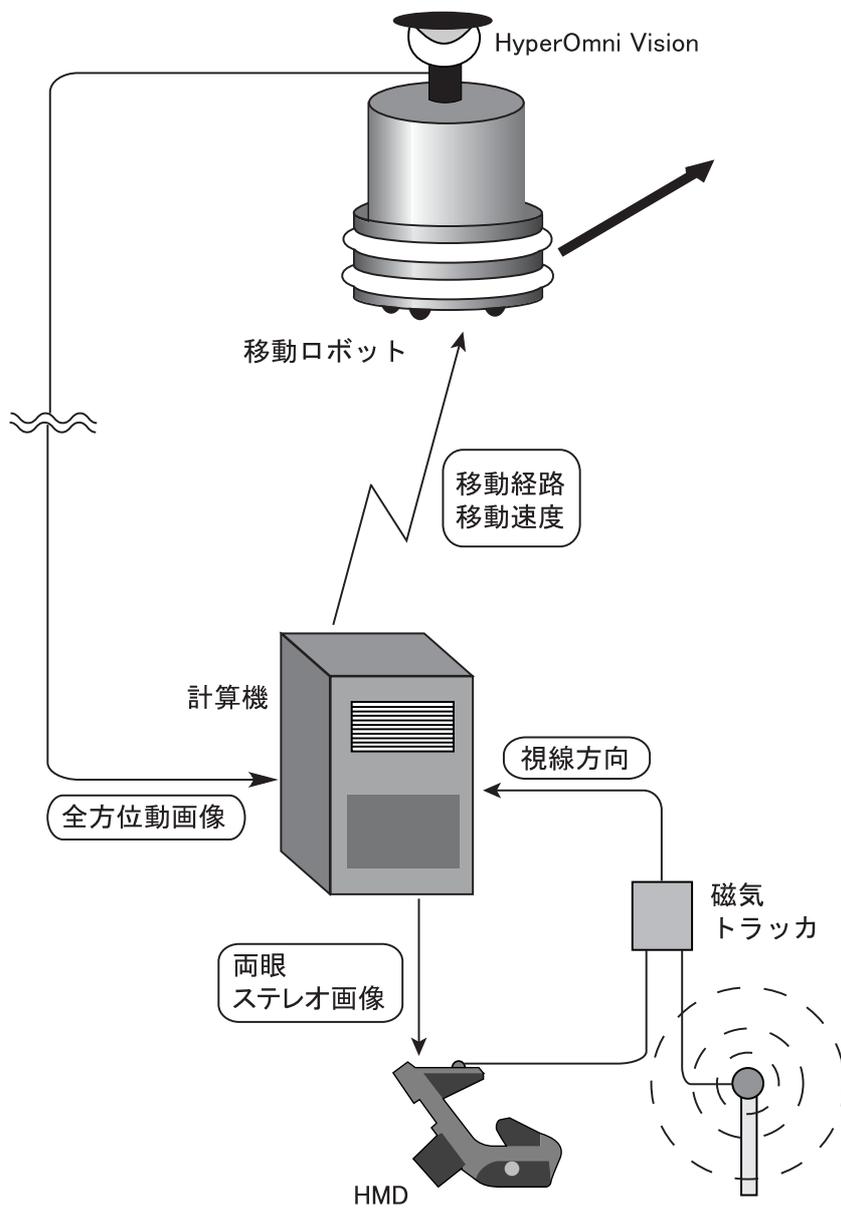


図 20 1 入力型システム構成

を計測する．

3. 視線方向に応じた両眼ステレオ画像を生成し，HMD に表示する．

なお、計算機に入力される全方位動画の解像度は 720×486 であるが、全方位画像が写る領域だけを切り出し、 500×450 の解像度で保持する．この画像の切り出しには 8CPU で並列処理を行い、それ以外では 1CPU で処理を行っている．

このシステムでは、全方位動画と移動経路、速度の情報を記録することにより蓄積型のテレプレゼンスも行うことができる．蓄積型のテレプレゼンスでは、全方位動画をビデオに録画した後、ビデオを再生して計算機に入力しながら、撮影時の移動経路と速度の情報をを用いて両眼ステレオ画像の生成と提示を行う．

5.1.3 実験

実際に 1 入力型システムで両眼ステレオ画像の生成と提示の実験を行った．実験では、HyperOmni Vision を載せた移動ロボットは 17.5cm/s の一定速度で直進させた．計算機には最新の 2.03 秒間の全方位動画を保持させ、その動画の中間点における両眼ステレオ画像を生成する．つまり、計算機に保持されている全方位画像は、伝送されてくる全方位画像の最新の 61 枚であり、両眼ステレオ画像を生成する地点からは前後 30 枚の全方位画像が計算機に存在している．そのため、利用者に提示する両眼ステレオ画像は約 1 秒前の地点でのものとなる．また左右の画像は、全方位動画の撮影位置から物体までの仮定距離を 3m として任意視点画像を生成する．なお、生成する両眼ステレオ画像の眼間距離は 7cm、水平画角は 50 度、解像度は 640×480 画素である．

図 21 は、約 30 秒間の入力全方位動画の 5 秒間隔の全方位画像を示している．実験によって生成された両眼ステレオ画像を図 22 に示す．図 22 の (a) から (f) のそれぞれにおいて左が左眼画像、右が右眼画像であり、(a) から (f) へと時間が経過している．図 22 に示すそれぞれの両眼ステレオ画像の視線方向は、進行方向に対する水平方向の角度と垂直方向の角度である．

図 22 のそれぞれの左右の画像から視差があることが確認できる．図 22 の (b) の椅子、(d) の鞆、(f) の三脚という近い位置に存在する物体は、それよりも遠くにある物体と比較して視差が大きくなっていることが分かる．

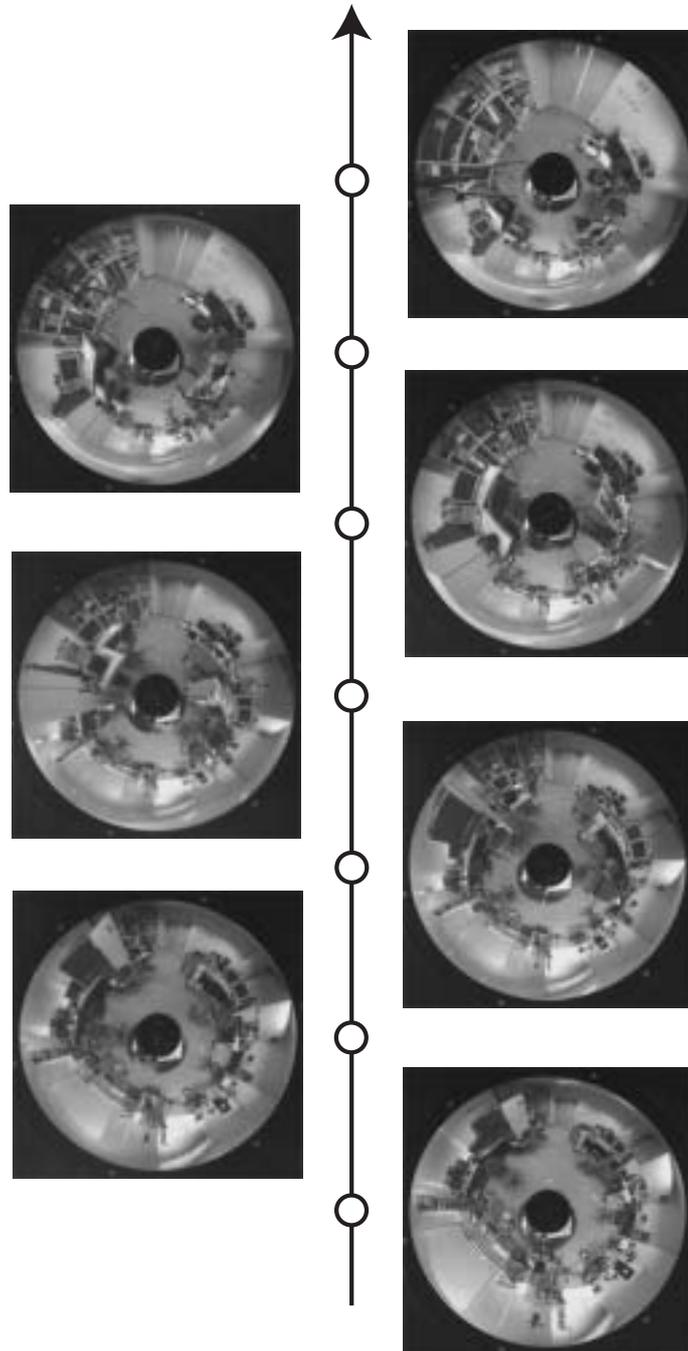


图 21 入力全方位動画像

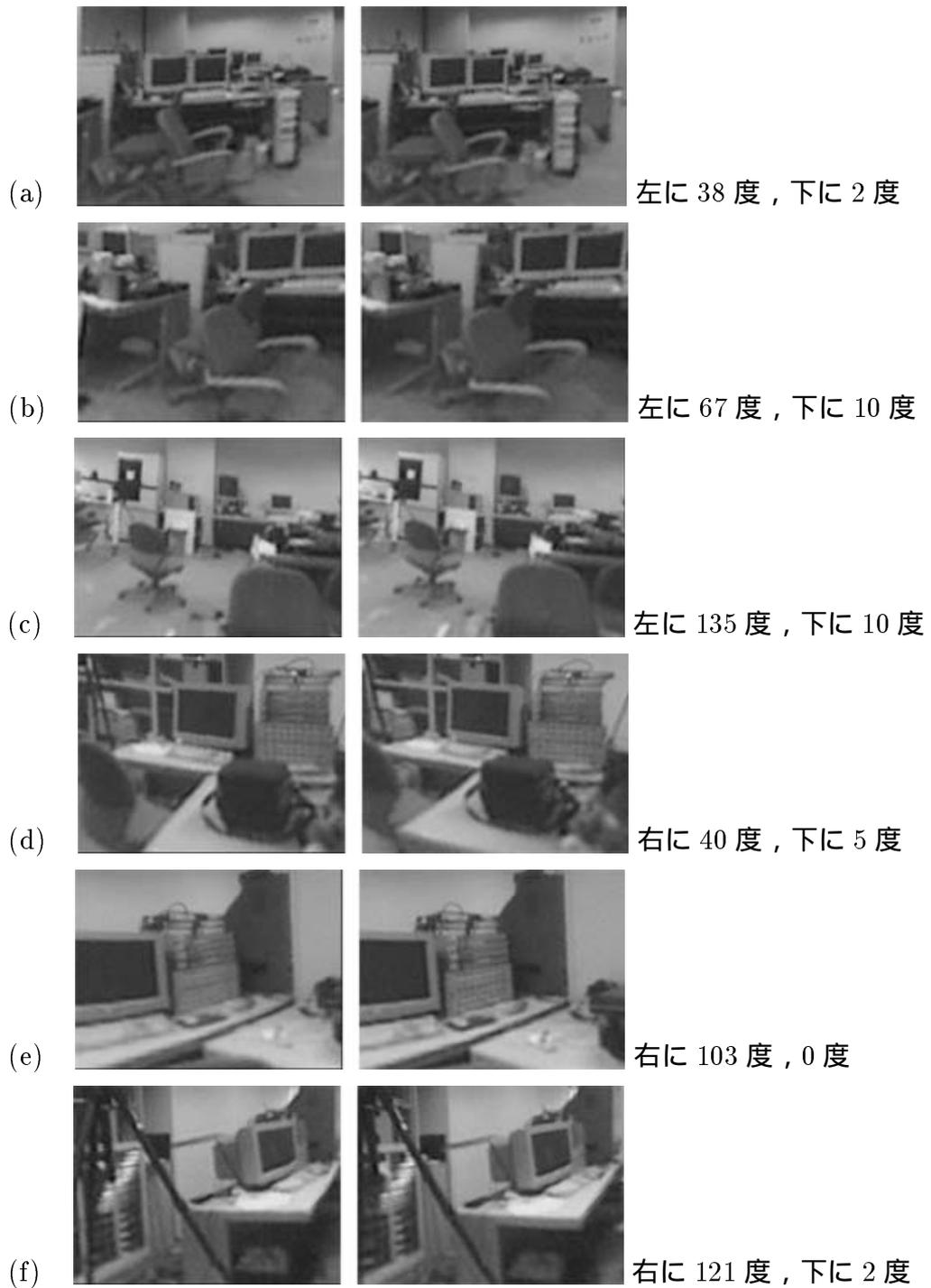


図 22 1 入力型システムにおいて生成された両眼ステレオ画像

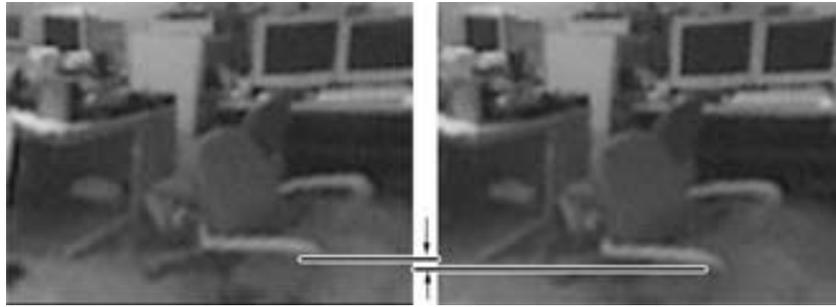


図 23 生成ステレオ画像に生じる縦視差

図 23に示す両眼ステレオ画像では、椅子のひじかけの位置が左右画像で縦にずれている。つまり左右画像の縦方向に関しては擬似的に生成しているため、左右の画像間には水平方向の視差だけでなく、縦方向の視差も生じている。しかし、実験では縦方向の視差が立体視に影響を与えるほどではなく、利用者はHMDで提示される両眼ステレオ画像により両眼立体視を行うことが可能であった。

両眼ステレオ画像の生成に要する時間は4.0ミリ秒で、視線方向が決定してから両眼ステレオ画像が提示されるまでの時間は37～70ミリ秒であった。HMDに提示される画像の更新速度は1/30秒であり、ビデオレートで両眼ステレオ画像を提示することが可能であった。全方位画像センサが17.5cm/sの一定速度で移動した2秒間の全方位動画像を計算機に常時保持し画像生成を行なった場合、保持された動画像の中間点で視線方向が進行方向に対して左右に36度から144度の両眼ステレオ画像を生成できた。図24に視線方向と生成に必要な全方位画像の範囲の関係を示す。図24の横軸は視線方向が進行方向に対して左または右になす角度で、縦軸は全方位画像の位置で、視点中心から進行方向あるいは逆方向に何枚目であるかを表し、左右画像それぞれの生成には図24の色を付けた範囲の全方位画像が必要であることを示している。図24から分かるように、視点中心から近い位置にある全方位画像は生成に用いないため、生成に用いる画像枚数は最大で24枚となる(視線方向が36度、144度の場合)。視点中心が経路上にあるとき、左右画像で用いる全方位画像の範囲は視点中心に対して対称になる。図24から、視線方向が36度より小さくなったとき、または144度より大きくなったときに

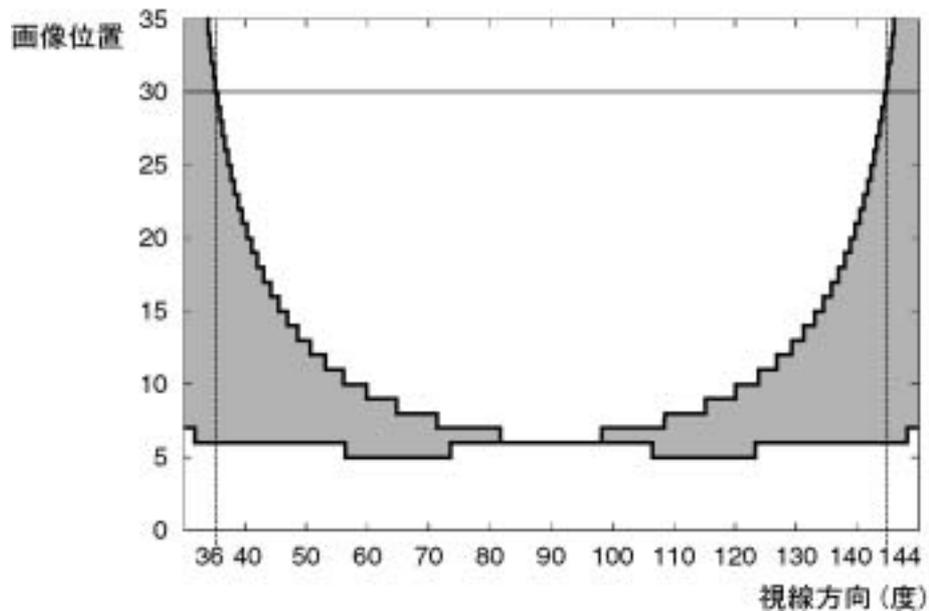


図 24 視線方向と生成に必要な画像の範囲の関係 (1 入力型システム)

は視点中心から 30 枚以上離れた位置の全方位画像が必要となることが分かる。

5.1.4 考察

直線経路の全方位動画像 1 つを用いて両眼ステレオ画像を生成し、提示する 1 入力型のテレプレゼンスシステムを作成し、実験を行った結果、両眼ステレオ画像の生成に要する時間は 4 ミリ秒、見直しに関する時間遅延は 37 ~ 70 ミリ秒という短時間であり、HMD に表示される両眼ステレオ画像のビデオレートでの更新が可能であった。また、利用者には 1.015 秒前の地点から見た画像が提示された。両眼ステレオ画像の生成に要する時間は 4 ミリ秒という短時間であったため、複数の利用者にそれぞれ異なる視線方向の両眼ステレオ画像を同時に生成・提示する、あるいは全周型のスクリーンなどにステレオ表示することをビデオレートで行うことも可能であると考えられる。

提案手法では、全方位動画像の撮影位置から物体までの距離を仮定して、任意視点画像生成手法により左右のステレオ画像を生成しているが、実際の物体まで

の距離とのずれにより両眼ステレオ画像に縦視差が生じた。実験では、立体視が行えなくなるような歪みは生じなかったが、未知の環境では物体までの適当な仮定距離を決定する必要があるが出てくる可能性がある。そこで、移動経路と速度が既知であることを用いて、全方位画像から物体までの距離を計算し、撮影環境に応じて仮定距離を変更して歪みを小さくすることも考えられる。

5.2 2 入力型蓄積テレプレゼンスシステム

1 入力型システムでは、撮影経路に平行な視線方向の両眼ステレオ画像の生成を行うことが困難である。そこで複数の全方位動画像を用いて、全方位の両眼ステレオ画像を生成することを考える。直交する2経路の全方位動画像から全方位の両眼ステレオ画像を生成することは可能であるが、全方位動画像をすべて計算機に蓄積する必要がある。そこで、2入力型システムでは眼間距離だけ離れ平行な経路をとる2つの全方位動画像を用いる。このシステムでは2つの全方位動画像を同時に計算機に伝送し、全方位の両眼ステレオ画像を生成することができる。全方位動画像は2回の撮影を行うことにより取得したため、2つの全方位動画像の取得後に動画像を再生して両眼ステレオ画像の生成・提示を行う蓄積型テレプレゼンスで行った。

5.2.1 両眼ステレオ画像生成

このシステムでは図25で示す経路1、経路2のような2つの経路の全方位動画像を用いて両眼ステレオ画像生成を行う。2つの全方位動画像の経路は平行で眼間距離だけ離れている。この2つの全方位動画像から、2つの経路の中線上に視点中心(両眼中心)が位置する両眼ステレオ画像を生成する。図25に示すように視点中心、視線方向が決まると、左眼、右眼の位置も決定する。図25のような位置関係にあるとき、左眼は経路2より経路1に近く、右眼は経路1より経路2に近い。そこで左眼画像は経路1の全方位動画像中の O_a から O_b 、右眼画像は経路2の全方位動画像中の O_c から O_d の全方位動画像を用いて生成する。両眼ステレオ画像の生成には、1入力型システムと同様に座標変換の計算量を減らす高速化手法[3]を用いる。

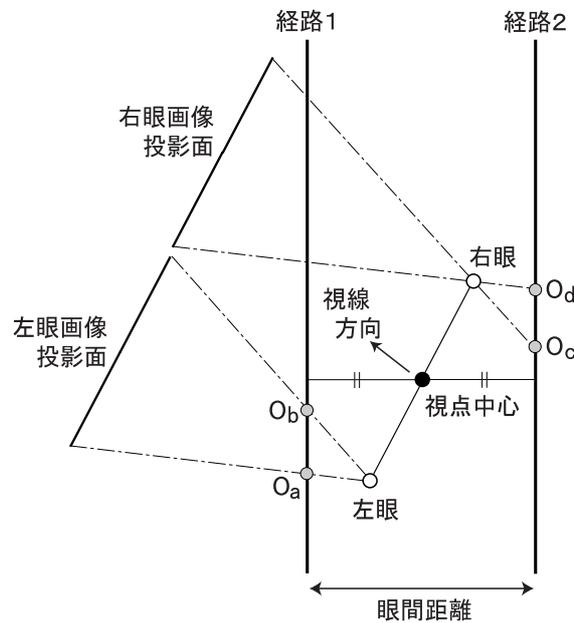


図 25 2 入力型システムにおける両眼ステレオ画像生成の図解

2つの全方位動画像を用いた両眼ステレオ画像を生成する手法では、図 26(a) に示すような進行方向と 90 度の角度をなす視線方向の場合には両目とも同じ経路から、眼間距離の半分の距離だけ離れた視点位置で生成する。図 26(a) では左眼は経路 1 の全方位動画像中の T_a から T_b 、右眼は経路 1 の全方位動画像中の T_c から T_d の全方位画像を用いて生成される。また、図 26(b) のように視線方向が進行方向に近づいた場合でも、左眼、右眼の位置はどちらかの経路に近くなるため必要となる全方位画像数は大きく増加しない。図 26(b) では、左眼は経路 1 の全方位動画像中の T_e から T_f 、右眼は経路 2 の全方位動画像中の T_g から T_h の全方位画像を用いて生成される。図 26(c) に示すように視線方向が進行方向と同じであるときには左眼、右眼の位置が経路上となるので、それぞれの目の位置に存在する全方位画像 1 枚から平面透視投影画像を作ることによって両眼ステレオ画像を生成できる。

視線方向を図 26 の (b) から (c) へ変化させていくと図 27 のような視線方向となる。図 27 の示す視線方向の左眼画像を生成する場合、左眼から P 点への光線は

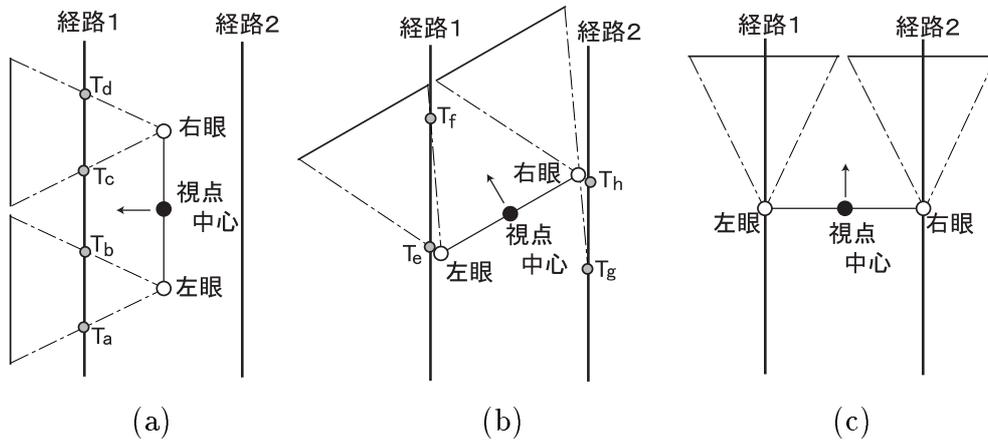


図 26 視線方向を変化させたときの両眼ステレオ画像生成

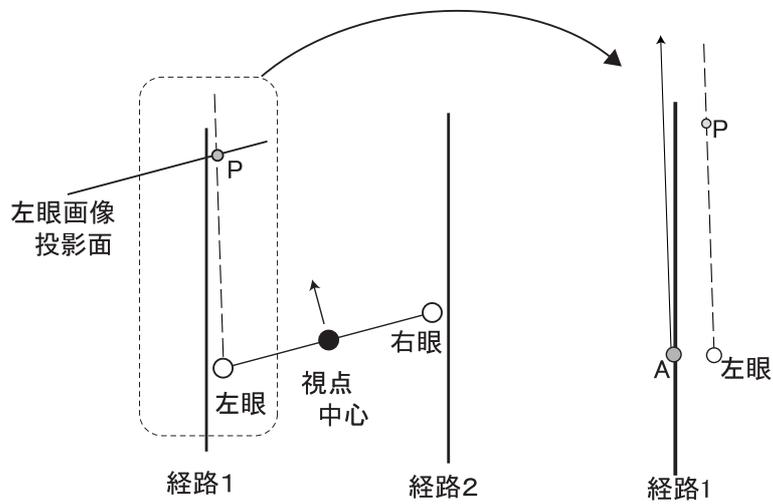


図 27 進行方向に近い光線が必要な場合の生成

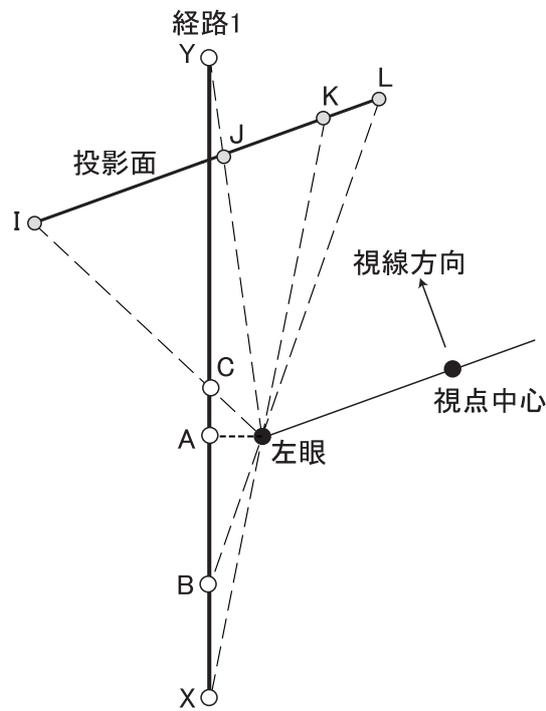


図 28 つなぎ目にずれの出る場合の生成

経路に平行な方向に近い場合、左眼画像の P 点を含む 1 列の生成には遠く離れた地点の全方位画像が必要となる。そこで必要な光線が経路と平行に近い、あるいは平行である場合には、目の位置に最も近い経路上の点の全方位画像を用いて生成を行う。つまり、図 27 の右に示すように左眼の位置に最も近い経路上の点 A の全方位画像から、左眼から P 点へ方向と平行な方向への画像を生成することで左眼画像の P 点を含む 1 列とする。この生成方法を用いると、左眼から P への光線と点 A からそれと平行に進む光線は異なるため、水平方向に関して擬似的な生成となる。この場合、水平方向に生じる歪みは物体が近くにあれば大きくなり、遠くにあれば小さくなる。

以上の手法により図 28 のような視線方向の左眼画像を生成することを考える。経路 1 の全方位動画は点 X から点 Y までが存在するとする。生成手法では、左眼画像の点 I から点 J までは点 C から点 Y までの全方位画像から生成し、点 J

から点 K までは存在しない全方位画像が必要となるため、経路上で左眼に最も近い点 A の全方位画像から擬似的に生成する。最後に、点 K から点 L までは点 X から点 B までの全方位画像を用いて生成する。このとき、左眼画像の点 I から点 J までの部分と点 K から点 L までの部分では生じる縦方向の歪みは逆になる。さらに、点 J から点 K までは点 A の全方位画像 1 枚から生成するため縦方向の歪みは生じない。したがって、左眼画像の点 J の列と点 K の列において、縦方向の歪みの違いからつなぎ目にずれが生じる。

5.2.2 システム構成

2 入力型システムでは眼間距離だけ離れて平行に進む 2 つの全方位動画像を必要とするため、2 つの全方位動画像をビデオに記録してから、両眼ステレオ画像の生成と提示を行う。2 入力型システムの構成は図 29 に示す。システムで使用した機器は 1 入力型システムと同様で表 2 の通りである。2 入力型システムは蓄積型テレプレゼンスシステムであるため、図 29 に示すように全方位動画像の取得と両眼ステレオ画像の生成・提示の 2 つの段階に分けられる。

2 つの全方位動画像の取得は、全方位画像センサ HyperOmniVision を載せた移動ロボットを用いて行う。移動ロボットは一定速度で、眼間距離だけ離れ平行な 2 つの直線経路を移動し、全方位画像センサより得られる 2 経路の全方位動画像をビデオで記録する。また、ロボットの移動経路と速度の情報はファイルとして保存する。

両眼ステレオ画像の生成と提示は以下の流れで処理を行う。

1. 2 つの全方位動画像を録画したビデオを同時に計算機に伝送し、計算機はそれぞれの全方位動画像の一定時間を保持する。
2. 利用者の装着する HMD に取り付けられた 3D 磁気トラッカで視線(頭部の向き)を計測する。
3. 全方位動画像の移動経路と速度の情報をを用いて、視線方向に応じた両眼ステレオ画像を生成し、HMD に表示する。

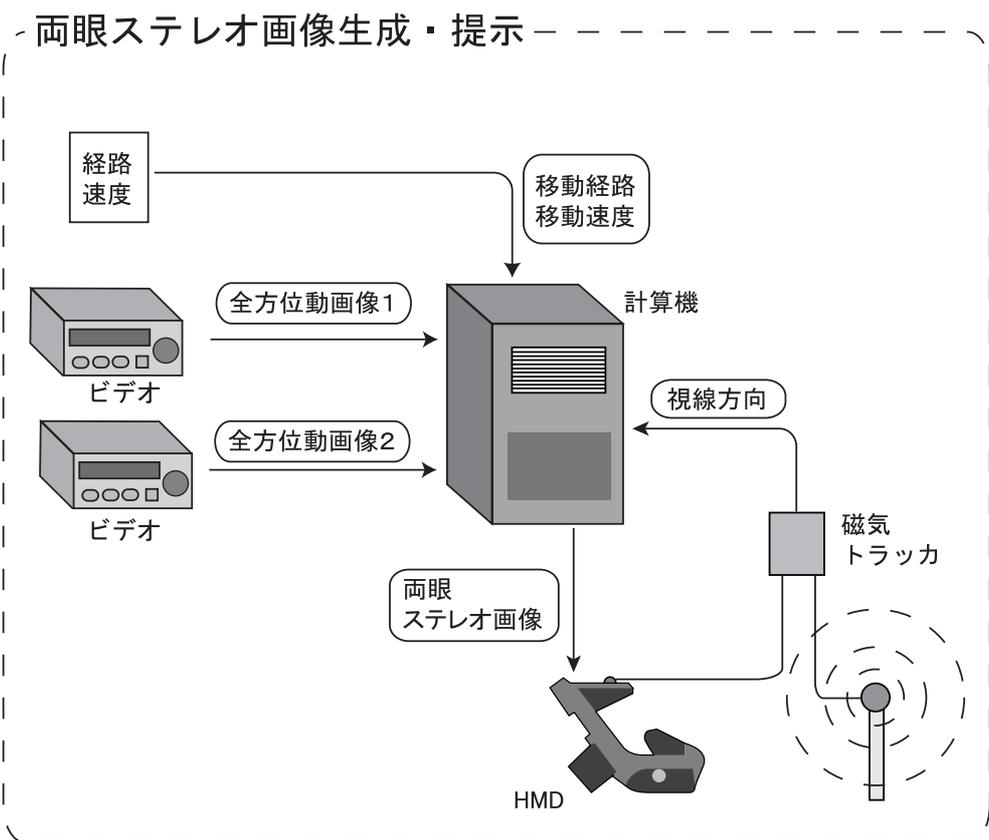
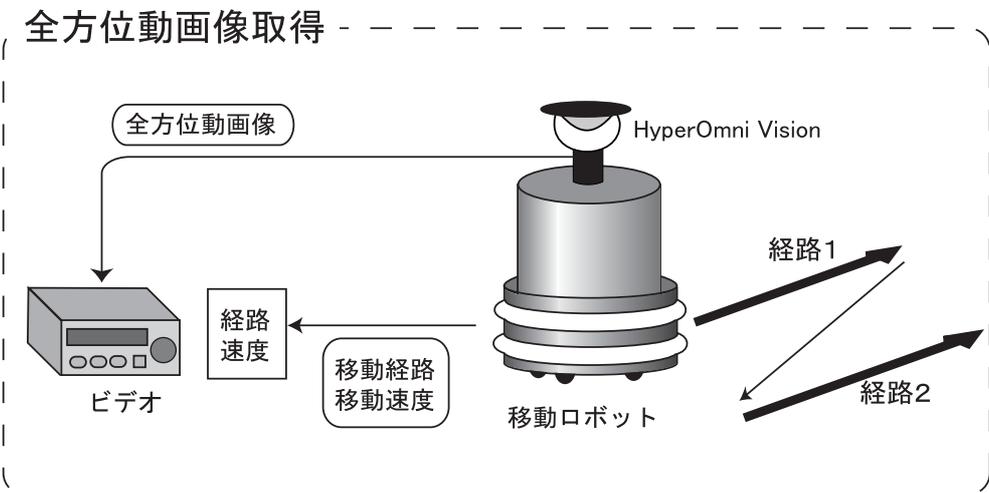


図 29 2 入力型システム構成

また、1入力型システムと同様に計算機には全方位画像の必要領域だけを切り出して保持する。2入力型システムでは、2つの全方位動画像を計算機に入力しているため、ビデオレートで表示を行うためには2枚の全方位画像の必要領域切り出しと切り出した画像のテクスチャメモリへの格納を1/30秒以内に行わなければならないが、現在のシステムにおいてはこの処理を1/30秒以内に行うことができない。そこで本システムでは、2つの全方位動画像の入力画像をそれぞれ1枚おきに読み込み、保持する。つまり、1秒間にそれぞれの全方位動画像で15枚の全方位画像が保持される。

5.2.3 実験

2入力型システムによる両眼ステレオ画像の生成、提示の実験を行った。2つの全方位動画像の取得はHyperOmni Visionを載せた移動ロボットを8.75cm/sの一定速度で、平行で眼間距離だけ離れた2つの経路を移動させることにより行う。計算機はビデオから伝送される2つの全方位動画像それぞれの最新2.2秒間を保持する。生成する両眼ステレオ画像の視点中心の位置は、計算機に保持している2つの全方位動画像の中間地点の midpoint とする。最新2.2秒間の全方位動画像を保持することとしているが、システムでは1枚おきに画像が読み込まれるため、計算機に保持される全方位画像枚数はそれぞれの経路で33枚となる。計算機には最新2.2秒間のそれぞれの全方位動画像を蓄積しているため、提示する両眼ステレオ画像はビデオで再生された2つの全方位動画像の1.1秒前の地点の midpoint でのものとなる。なお、生成する両眼ステレオ画像の眼間距離は7cm、水平画角は50度、解像度は640×480画素である。

図30に入力された2経路の全方位動画像を示す。図30は2経路の全方位動画像の4秒おきの画像で、全方位動画像は2経路それぞれ約12秒間である。実験で生成された両眼ステレオ画像を図31に示す。図31のそれぞれのステレオ画像において左側が左眼画像、右側が右眼画像であり、(a)から(f)へ時間が経過している。図31に示すそれぞれの両眼ステレオ画像の視線方向は、進行方向に対する水平方向の角度と垂直方向の角度である。

図31の(a)～(f)から視差のある両眼ステレオ画像が生成されていることが確

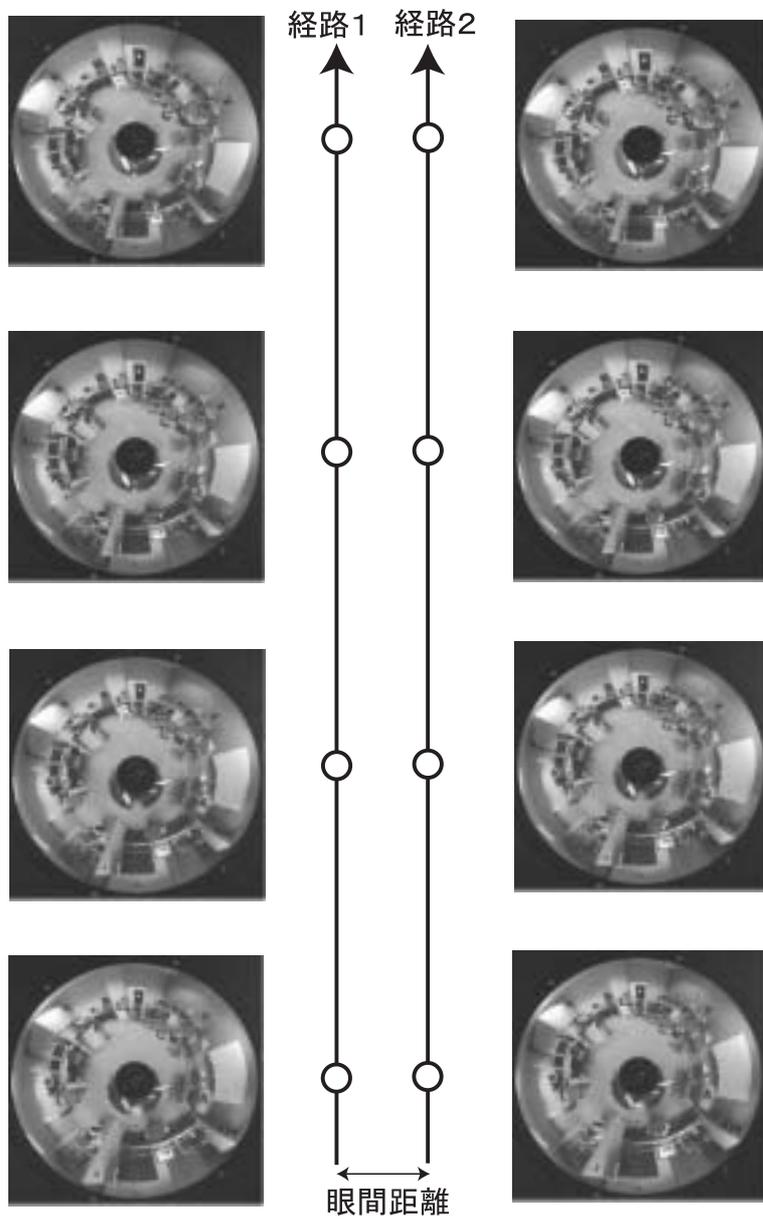


図 30 2 経路の入力全方位動画像

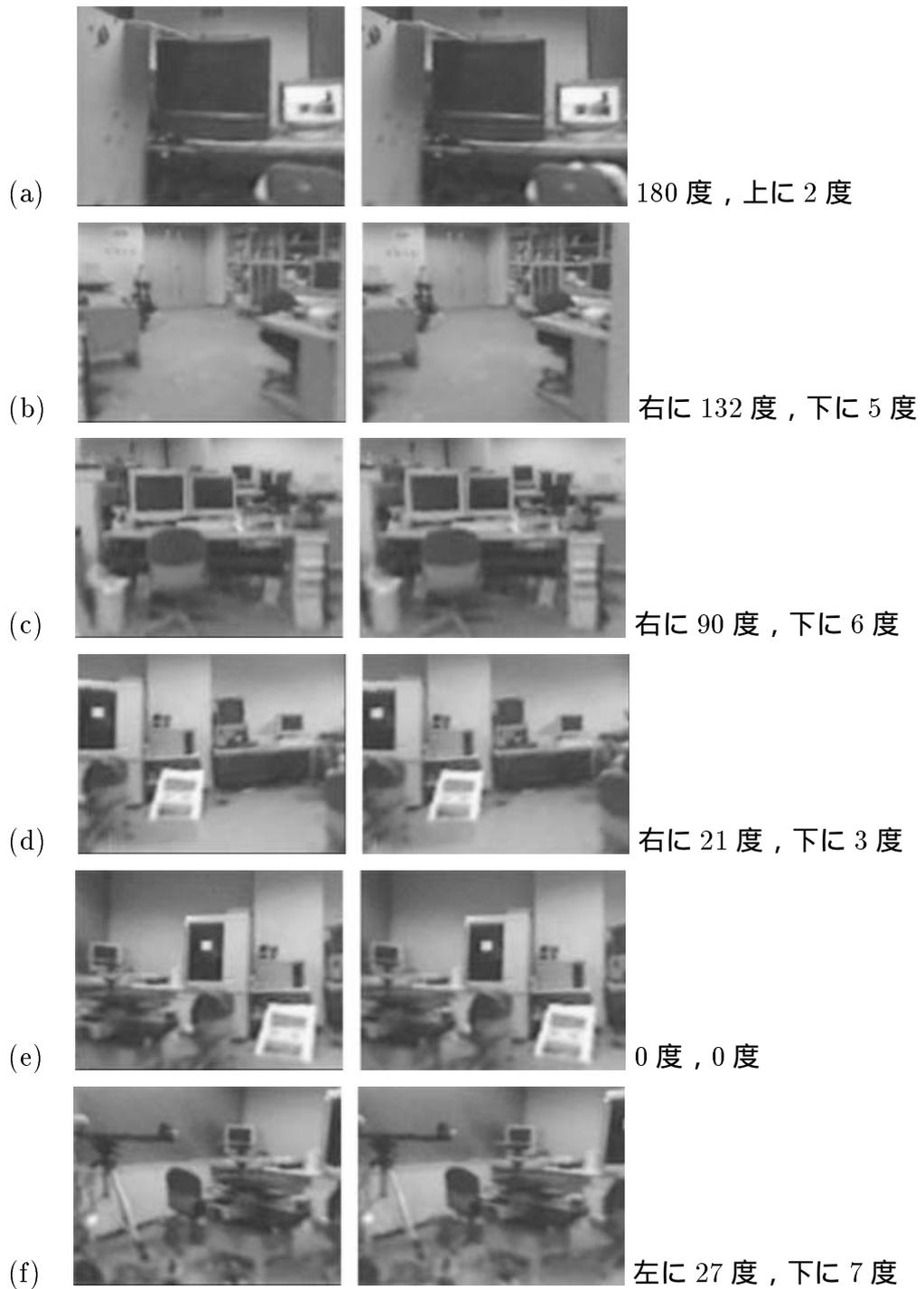


図 31 2 入力型システムにおいて生成された両眼ステレオ画像



図 32 生成ステレオ画像に生じる縦方向のずれ

認できる．2 入力型システムでは，(d)，(f) のように 1 入力型システムでは生成できなかった視線方向が進行方向に近い両眼ステレオ画像や，(a)，(e) のように進行方向あるいは進行方向の逆方向を向いたときの両眼ステレオ画像が生成でき，全方向で両眼ステレオ画像を生成することが可能であった．また，(c) は進行方向に対して真横を向いた時の両眼ステレオ画像で，左右画像とも同じ経路の全方位動画像から生成した．

図 32 は，図 31(d) の右眼画像である．視線方向が進行方向に近いときの両眼ステレオ画像であるため，図 32 の画像の生成では目の位置に最も近い全方位画像を用いて，水平方向に関して擬似的に生成する部分ができるが，目の位置と経路上で目に最も近い位置の全方位画像の距離は短く，物体までの距離も長いため，水平方向にはほとんどずれは生じていない．しかし，前述したように縦方向の歪みが増加することにより，図 32 の矢印で示す列で縦方向にずれていることが確認できる．

両眼ステレオ画像の生成に要する時間は 4.3 ミリ秒，視線方向が決定してから両眼ステレオ画像が提示されるまでの時間は 70.3 ~ 136.3 秒であり，HMD に提示される画像の更新間隔は 1/15 秒であった．HMD の更新間隔が 1/15 秒であるのは，入力全方位画像の切り出しとテクスチャメモリへの格納に時間を要するためであり，両眼ステレオ画像の生成は実時間で行うことができた．

実験では，それぞれの経路で 33 枚の全方位画像を保持して生成を行った．図 33 の上に示すように，経路 1 で保持する全方位動画像の中間点を 0 として，33 枚の全方位画像に -16 から 16 までの番号をつける．このとき，視線方向と生成に用いる画像の範囲の関係を図 33 のグラフに示す．グラフの横軸は視線方向が進行方向に対して左になす角度で，グラフでは左眼画像の生成に用いる画像の範囲を色を付けて表している．視線方向が -27 度から 27 度の範囲で 0 度以外の場合には，計算機に保持されていない全方位画像が必要となるため，目の位置に最も近い全方位画像を用いて擬似的に生成を行った．疑似生成のために用いた目に最も近い全方位画像の位置を図 33 の点線で示す．また，視線方向が進行方向に対して左に 90 度から 180 度，-90 度から -180 度の場合は，図 33 において -90 ~ 0 ~ 90 である横軸の目盛りを -90 ~ 180 ~ 90 とし，縦軸を経路 2 での画像範囲としたものと同じである．右眼の場合は，左眼のグラフを上下反転させたものとなる．

5.2.4 考察

2 つの直線経路の全方位動画像を用いて両眼ステレオ画像を生成し，提示する 2 入力型のテレプレゼンスシステムについて述べた．このシステムでは，記録された 2 つの全方位動画像を入力することにより，全方向の両眼ステレオ画像を生成できる．実験において，両眼ステレオ画像の生成に要する時間は 4.3 ミリ秒，見回しに関する時間遅延は 70.3 ~ 136.3 秒であり，提示画像の更新間隔は 1/15 秒であった．また，利用者には 1.1 秒前の地点から見た両眼ステレオ画像が提示された．視線方向が進行方向に近いとき，生成される画像にずれが生じた．これは，必要な全方位画像が計算機内に存在しないために目の位置に最も近い全方位画像から生成を行う画像内の領域と，その両側の領域で縦方向の歪みが増加することに原因がある．進行方向付近に写る物体までの距離を計算することにより，

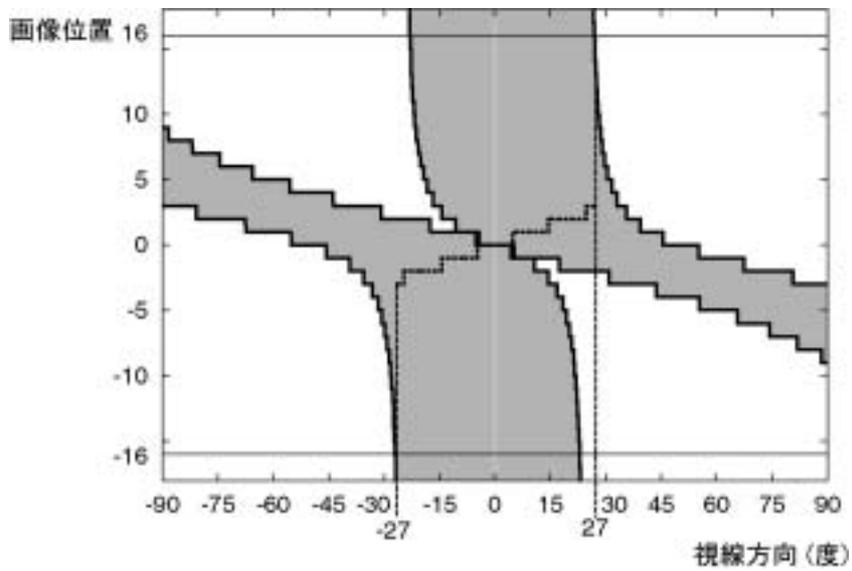
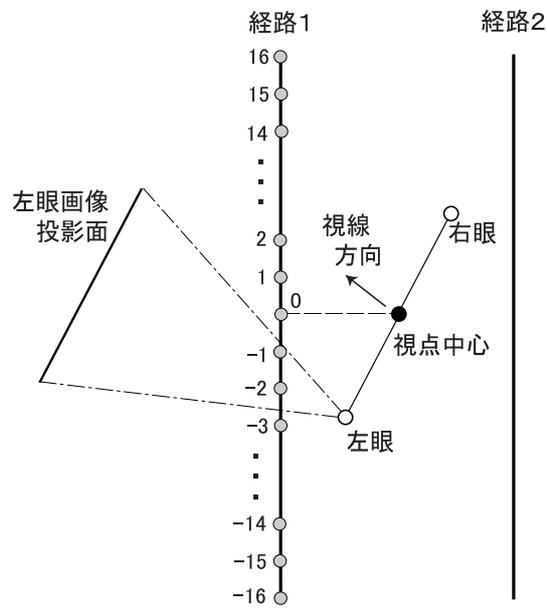


図 33 視線方向と生成に必要な画像の範囲の関係 (2 入力型システム)

縦方向に歪みをなくせば，このずれもなくすることができると考えられる．

現在のシステムでは，あらかじめ2つの経路の全方位動画像を撮影する必要があり，撮影は2つの経路を順番に移動ロボットを移動させることにより行う．移動ロボットにはモータの回転センサがあり，その位置と姿勢を検出することができるが，床の凹凸や回転センサの精度などの問題から位置，姿勢には誤差が生じる．そのため，長い経路を移動させると2つの経路間の関係にもずれが生じ，両眼ステレオ画像の生成ができなくなる．そこで，全方位画像センサを眼間距離で並べて同時に2経路の全方位動画像を取得することが考えられる．全方位画像センサを並べると，取得される全方位画像にもう一方の全方位画像センサが写り込むことになるが，その部分が広域にわたらないのであれば，写り込んでいる部分を用いることなく両眼ステレオ画像を生成することも可能である．本研究で使用した HyperOmni Vision ver.2A は鏡直径が 150mm であるため眼間距離で並べることはできないが，HyperOmni Vision と同原理の小型の全方位画像センサを用いると眼間距離で並べることができ，2経路の全方位動画像を同時に取得できる．2経路の全方位動画像が同時に取得できれば，2入力型システムにおいても実時間テレプレゼンスを行うことが可能となる．

6. おわりに

本論文では，撮影位置が既知である全方位画像センサを用いて取得した全方位画像系列から，撮影位置付近の任意視点画像を擬似的に生成する手法を提案した．また，その手法を用いることにより，見回しと立体視が可能なテレプレゼンスシステムの方式を提案し，実際にプロトタイプシステムの開発を通して，その有効性を確認するとともに問題点を明らかにした．任意視点画像生成の提案手法では，全方位画像センサ HyperOmni Vision を水平に直線移動させることにより取得した全方位画像系列から，光線情報を用いて撮影経路を含む水平面に位置する任意視点を生成する．この提案手法を用いると，直交する2経路の全方位動画画像から，経路付近の任意の視点位置で任意方向の画像を生成することが可能であることを実験から確認した．また，見回しと立体視が可能なテレプレゼンスシステムは，1つの全方位動画画像を用いる1入力型システムと2つの全方位動画画像を用いる2入力型システムをを作成し，実験を行った．1入力型システムでは，必要な画像枚数が増加することから視線方向が制限されるが，実時間テレプレゼンスを行うことが可能である．2入力型システムでは，2つの全方位動画画像の取得後に両眼ステレオ画像の生成・提示を行う蓄積型テレプレゼンスとなるが，全方位の両眼ステレオ画像を生成することができる．2つのシステムでは，利用者は見回しと立体視を行うことができ，臨場感豊かな遠隔地の情景提示が可能である．また，両眼ステレオ画像は約4ミリ秒という短時間で生成されるため，見回しに関して時間遅延が少ない．

提案手法では少ないデータ量で高速に画像の生成を行うために生成画像の縦方向に歪みが発生し，両眼ステレオ画像に縦視差やずれが生じる．そこで今後の課題として，全方位動画画像から物体の距離を測定し，縦方向の歪みを軽減する生成手法の検討が挙げられる．また，小型の全方位画像センサを眼間距離で並べて移動させ，2経路の全方位動画画像を同時に取得することで，実時間で全方位の両眼ステレオ画像提示が行えるシステムの構築が考えられる．

謝辞

本研究を進めるにあたり，終始暖かく御指導を頂いたソフトウェア基礎講座 横矢 直和 教授に厚く御礼申し上げます．また，本学での研究におけるあらゆる面で，多大な御助言，御協力を賜りましたことを深く感謝いたします．

副指導教官として御助言を頂いたロボティクス講座 小笠原 司 教授，ならびにソフトウェア基礎講座 竹村 治雄 助教授に深く感謝致します．

本研究への御助言や研究環境の整備に御協力を頂いたソフトウェア基礎講座 山澤 一誠 助手ならびに岩佐 英彦 助手に深く感謝いたします．

多くの御助言を頂いたソフトウェア基礎講座 佐藤 哲 氏，神原 誠之 氏，松宮 雅俊 氏に深く感謝いたします．

二年間様々な面で御協力を頂いた同輩方，一年間ではありますがお世話になった後輩方に心より感謝申し上げます．

また，日々の研究活動を支えてくださったソフトウェア基礎講座 事務補佐員 福永 博美女史に感謝いたします．

参考文献

- [1] Saied Moezzi, Ed.: “Special issue on immersive telepresence”, *IEEE Multi-Media*, vol. 4, no. 1, pp. 17-56, 1997.
- [2] 景観放送サービス (LiveCast)
“<http://biz.biglobe.ne.jp/service/livecast/index.html>”
- [3] Y. Onoe, K. Yamazawa, H. Takemura, and N. Yokoya: “Telepresence by real-time viewpoint image generation from omnidirectional video streams”, *Computer Vision and Image Understanding*, vol. 71, no. 2, pp. 154-165, 1998.
- [4] Y. Yagi: “Omnidirectional sensing and its applications”, *IEICE Trans. on Information and Systems*, vol. E82-D, no. 3, pp. 568-579, 1999.
- [5] S. E. Chen: “QuickTime VR - An image-based approach to virtual environment navigation”, *Proc. SIGGRAPH'95*, pp. 29-38, 1995.
- [6] M. Hirose, K. Yokoyama, and S. Sato: “Transmission of realistic sensation: Development of a virtual dome”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 23-28, 1992.
- [7] 中尾, 柏谷, 兼吉: “回転鏡を用いた広視野カメラの試作”, *情報処理学会研究報*, CVIM115-16, 1999.
- [8] Z. L. Cao, S. J. Oh, and E. L. Hall: “Dynamic omnidirectional vision for mobile robots”, *Proc. SPIE Intelligent Robots and Computer Vision*, vol. 579, pp. 405-414, 1985.
- [9] J. Hong, X. Tan, B. Pinette, R. Weiss, and E. M. Riseman: “Image-based homing”, *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 620-625, 1991.
- [10] Y. Yagi, S. Kawato, and S. Tsuji: “Real-time omnidirectional image sensor (COPIS) for vision-guided navigation”, *IEEE Trans. Robotics and Automation*, vol. 10, no. 1, pp. 11-22, 1994.

- [11] 山澤, 八木, 谷内田: “移動ロボットのナビゲーションのための全方位視覚センサ HyperOmni Vision の提案”, 電子情報通信学会論文誌 (D-II), J79-D-II, no. 5, pp. 698–707, 1996.
- [12] S. K. Nayar: “Catadioptric omnidirectional camera”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 482-488, 1997.
- [13] T. Kawanishi, K. Yamazawa, H. Iwasa, H. Takemura, and N. Yokoya: “Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors”, *Proc. Int. Conf. on Pattern Recognition*, vol. I, pp. 485-489, 1998.
- [14] V. N. Peri and S. K. Nayar: “Generation of perspective and panoramic video from omnidirectional video”, *Proc. Image Understanding Workshop*, pp. 243-245, 1997.
- [15] S. E. Chen and L. Williams: “View interpolation for image synthesis”, *Proc. SIGGRAPH'93*, pp. 279-288, 1993.
- [16] M. Kato and H. Noyama: “Interactive visual simulation in a quasi-three-dimensional world based on the structuralization of images”, *Proc. Int. Conf. on Artificial Reality and Tele-Existence*, pp. 101-110, 1995.
- [17] S. M. Seitz and C. R. Dyer: “View Morphing”, *Proc. SIGGRAPH'96*, pp. 21-30, 1996.
- [18] J. I. Park and S. Inoue: “Image-based rendering from multi-view images”, *Journal of ITE*, vol. 52, no. 3, pp. 371-376, 1998.
- [19] E. H. Adelson and J. Bergen: “The plenoptic function and the elements of early vision”, *Computational Models of Visual Processing*, Chapter 1, M. Landy and J. A. Movshon, Ed., MIT Press, 1991.
- [20] L. McMillan and G. Bishop: “Plenoptic modeling: An image-based rendering system”, *Proc. SIGGRAPH'95*, pp. 39-46, 1995.

- [21] M. Levoy and P. Hanrahan: “Light field rendering”, *Proc. SIGGRAPH'96*, pp. 31-42, 1996.
- [22] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen: “The lumigraph”, *Proc. SIGGRAPH'96*, pp. 43-54, 1996.
- [23] 藤井, 木藤, 金子, 原島: “3次元統合画像通信を目指したホログラムと光線空間の相互変換”, 3次元画像コンファレンス'95, pp. 220-225, 1995.
- [24] H. Y. Shum and L. W. He: “Rendering with concentric mosaics”, *Proc. SIGGRAPH'99*, pp. 299-306, 1999.
- [25] S. Peleg and J. Herman: “Panoramic mosaics by manifold projection”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 338-343, 1997.
- [26] 山口, 山澤, 竹村, 横矢: “全方位動画像を用いた任意方向の両眼ステレオ画像の生成”, 情報処理学会第59回全国大会, 2-245, 1999.
- [27] 山口, 山澤, 竹村, 横矢: “全方位動画像を用いた任意方向両眼ステレオ画像の実時間生成と提示”, 電子情報通信学会技術研究報告, PRMU99-159, 1999.
- [28] 高橋, 川崎, 池内, 坂内: “全方位画像によるレンダリング手法の提案”, 情報処理学会研究報, CVIM119-5, 1999.
- [29] J. Neider, T. Davis, and M. Woo: *OpenGL Programming Guide*, Addison-Wesley Publishing Company, 1993.