

修士論文

全方位画像センサを用いたネットワーク対応型の  
テレプレゼンスシステムと遠隔監視システム

森田 真司

2003年2月7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。

森田 真司

審査委員： 横矢 直和 教授  
砂原 秀樹 教授  
山澤 一誠 助教授

# 全方位画像センサを用いたネットワーク対応型の テレプレゼンスシステムと遠隔監視システム\*

森田 真司

## 内容梗概

テレプレゼンスや遠隔監視は、バーチャルツアーやロボットの遠隔操縦、遠隔共同作業などさまざまな分野で利用が可能である。テレプレゼンスや遠隔監視では、ネットワークを用いて、遠隔地の様子を広範囲に、かつ臨場感豊かに伝送することが重要となる。従来のテレプレゼンスや遠隔監視では、回転カメラを用いる方法や全方位画像センサを用いる方法が提案されている。前者においては見回しに機械的な時間遅延が生じる。これに対し、後者においては観測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いが、全方位画像センサで撮影した全方位画像を直接、計算機に伝送しており、ネットワークを介した実装には至っていなかった。特に遠隔監視においては、ネットワークを用いなければ、ビル全体など多くのセンサを必要とする環境には適用できない。

そこで本研究では、ネットワークを介して全方位画像を伝送するテレプレゼンスシステムと遠隔監視システムを開発した。本テレプレゼンスシステムはネットワークを介して遠隔地の全方位画像を伝送し、その画像より利用者の視線方向の画像を実時間で生成し、提示する。本研究で開発したテレプレゼンスシステムを評価するために、有線ネットワークを用いた学内実験と有線・無線ネットワークを用いた学外実験を行った。共に 30fps で全方位画像を受信することができ、かつ時間遅延が少なく複数人が同時に異なる方向を見回しすることができるテレプレゼンスシステムの有効性を確認した。また、上述のネットワークを介したテレプレゼ

---

\* 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-MT0151110, 2003年2月7日.

ンスシステムを応用した遠隔監視システムは, 監視環境側をサーバ側, 監視者側をクライアント側としたサーバ/クライアントモデルである. まずサーバ側において監視したい環境を複数の全方位画像センサによって撮影し, その画像より移動物体を検出し, センサからの移動物体の方位を推定する. 次にその全方位画像と方位情報を, ネットワークを利用してクライアント側に伝送する. クライアント側では時系列の方位情報を用いて移動物体の存在領域を推定し, 監視者に対し物体方向の画像を提示することにより追跡を行う. 本研究で開発した遠隔監視システムは, サーバ側において環境を自由に動く複数の人物の方位情報を約 0.2 秒の更新間隔で推定でき, かつ静止し続ける物体を背景と判断し, 移動物体として検出しないことを確認した. またクライアント側においては人物の方位情報は約 0.2 秒ごとに更新され, それに基づいた移動物体の位置推定を行うことができた. さらに人物の方位情報は約 0.2 秒ごとに更新されるものの, 提示画像は約 0.05 秒ごとに更新でき, 遠隔監視システムとしての有効性を確認した.

#### キーワード

ネットワークテレプレゼンス, Digital Video Transport System, 遠隔監視システム, 全方位画像センサ, 移動物体追跡

# Networked Telepresence and Remote Surveillance Systems Using Omnidirectional Image Sensors\*

Shinji Morita

## Abstract

Telepresence and remote surveillance are widely utilized for virtual tour, telerobotics and so on. For telepresence and remote surveillance, it is important that the system displays a wide scene to a remote user with rich presence via network. In most conventional telepresence and remote surveillance systems, active cameras and omnidirectional video cameras are generally used in order to acquire an image of the remote site. The former often suffers from the time delay of changing the user's viewing direction and updating the displayed image. The latter is not implemented over network so that it cannot be applied to remote surveillance using a large number of distributed sensors.

This paper proposes new networked telepresence and surveillance systems. The proposed telepresence system transports an acquired omnidirectional video stream of a remote site to an observation site through the internet, and then generates a view-dependent perspective image. Two kinds of experiments have been carried out for evaluating the system. The first is an intramural experiment through a campus LAN and the second is an extramural experiment through a wired LAN and optical wireless LAN. Experiments have shown that the time delay of changing the user's viewing direction and updating the displayed image is

---

\* Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0151110, February 7, 2003.

sufficiently small and multiple users can simultaneously look around from a single viewpoint in a visualized dynamic real world in different directions via internet. Therefore, these results show that the present system is applicable to real-time telepresence in the situation where the real world to be seen is far from the observation site. The surveillance system, which is an application of the networked telepresence system, is based on a server/client model: the server computers, each of which is connected to an omnidirectional video camera, are placed in the surveillance area and the client computer is placed in a user side. The servers detect moving objects and estimate their directions from sensors. The client estimates object positions from their directions received from the servers and presents object-centered perspective images to the user like the telepresence system. In experiments, the servers successfully estimate multiple object directions every 0.2msec ignoring reposed objects. The client estimates object positions using those received directions and refreshes the image display every 0.05msec. These results show the possibility of real-time remote surveillance.

**Keywords:**

Networked Telepresence, Digital Video Transport System, Networked Remote Surveillance System, Omnidirectional Imaging, Motion Tracking

# 目次

<b>1. はじめに</b>	<b>1</b>
1.1 テレプレゼンスと遠隔監視に関する従来研究	2
1.2 本研究の方針と概要	6
<b>2. ネットワークテレプレゼンスシステム</b>	<b>8</b>
2.1 全方位画像センサ HyperOmni Vision	8
2.2 全方位画像の取得	9
2.3 全方位画像の伝送	11
2.4 透視投影画像の生成	12
2.5 ネットワークテレプレゼンスシステムの構築	13
2.5.1 学内実験とシステム構成	14
2.5.2 学内実験の結果	15
2.5.3 学外実験とシステム構成	15
2.5.4 学外実験の結果	18
2.5.5 考察	21
<b>3. 遠隔監視システム</b>	<b>22</b>
3.1 センサの初期設定	25
3.2 全方位画像からの注目物体の検出	26
3.2.1 背景更新を行う背景差分による注目物体の検出	26
3.2.2 注目物体の方位情報の導出	29
3.3 注目物体の追跡	30
3.3.1 多角形領域の辺導出による存在領域推定	30
3.3.2 注目物体方向の画像提示	37
3.4 遠隔監視システムの構築	37
3.4.1 システム構成	38
3.4.2 人物の検出・追跡実験	38
3.4.3 考察	49

4. おわりに	51
謝辞	54
参考文献	55
研究業績	59

## 目 次

1	カメラの向きを制御するテレプレゼンスシステム . . . . .	3
2	QuickTime VR . . . . .	3
3	Virtual Dome . . . . .	3
4	全方位画像センサを用いたテレプレゼンスシステム . . . . .	3
5	全方位画像センサと回転カメラを組み合わせた監視システム [19] . . . . .	5
6	HyperOmni Vision . . . . .	10
7	全方位画像と平面透視投影画像 . . . . .	10
8	Digital Video Transport System の処理概要 . . . . .	12
9	全方位画像と平面透視投影画像の関係 . . . . .	13
10	学内実験におけるネットワークテレプレゼンスシステムのシステム構成 . . . . .	16
11	学内実験における見回しの様子 . . . . .	16
12	学内実験の様子 . . . . .	17
13	学外実験におけるネットワークテレプレゼンスシステムのシステム構成 . . . . .	19
14	学外実験における見回しの様子 . . . . .	19
15	学外実験の様子 . . . . .	20
16	遠隔監視システムの構築図 . . . . .	22
17	ローカル座標系と世界座標系の角度のずれ . . . . .	25
18	極座標変換によるヒストグラムの作成 . . . . .	29
19	物体の方位角範囲 . . . . .	31
20	注目物体の存在領域 . . . . .	31
21	複数の人物の特殊な位置関係 . . . . .	32
22	直線の媒介変数表現 . . . . .	33
23	1つの物体の存在領域 . . . . .	35
24	物体1の信頼度 . . . . .	36
25	存在領域候補のクラス分け . . . . .	36
26	遠隔監視システムの構成 . . . . .	39

27	実験環境 . . . . .	41
28	監視者への提示画像 . . . . .	41
29	実験 1 で推定した存在領域の連続 CG . . . . .	42
30	実験 1 でのサーバにおける入力画像と注目物体の検出 . . . . .	43
31	実験 1 でのクライアントにおける提示画像 (図 28 を分割) . . . . .	44
32	実験 2 でのサーバにおける入力画像と注目物体の検出 . . . . .	45
33	実験 2 でのクライアントにおける提示画像 (図 28 を分割) . . . . .	46
34	実験 3 でのサーバにおける入力画像と注目物体の検出 . . . . .	47
35	実験 3 でのクライアントにおける提示画像 (図 28 を分割) . . . . .	48

## 表 目 次

1	学内実験におけるネットワークテレプレゼンスシステムの機器構成	14
2	学外実験におけるネットワークテレプレゼンスシステムの機器構成	18
3	遠隔監視システムの機器構成 . . . . .	39

# 1. はじめに

近年、バーチャルツアーや移動ロボットの遠隔操作、多人数が参加する遠隔通信会議、遠隔共同作業などが盛んに行われるようになってきており、これらの分野においては、テレプレゼンスや遠隔監視の技術が広く利用されている。テレプレゼンスとは、遠隔地の情景を提示し、その場にいるのと同様の感覚を与える仮想現実感技術[1]のことである。画像は音声や文字に比べて情報量が大きく、テレプレゼンスにおいて遠隔地の様子を広範囲に、かつ臨場感豊かに伝送することが望まれている。遠隔地に存在する移動ロボットを操縦する場合などでは、ロボットの周囲に障害物が存在する場合には、一方向の画像だけでは障害物全てを見ることができない。したがって、ロボットから送られた情報を基に広範囲かつ臨場感のある画像を提示し、遠隔地の自由かつ広範囲の見回しを行うことで、周囲を知覚しやすく、遠隔地での作業が効率良く行うことができ、操作性を向上させることが可能となる。

遠隔監視においても、テレプレゼンスと同様に広範囲の遠隔地の様子を取得し、複数の物体の同時検出および追跡を行う実時間監視が重要である。また、両システムともネットワークを介して遠隔画像を伝送することで、専用回線を設置する必要がなく、コストの削減が見込まれる。特に、遠隔監視においては、監視環境に存在する人物や物体の状況・動作を効率よく理解することができ、かつ複数の対象物体を視覚的に認識することが重要である。また、上記の要求事項を満たす遠隔監視システムは対象物体の自動検出による物体の移動経路の解析や移動ロボットの障害物回避などにも応用が可能である。

本研究では、ネットワークを介して遠隔地を自由に見渡すことができるテレプレゼンスシステムを開発した。また、テレプレゼンス技術の応用として、移動物体を自動検出及びその画像を提示する実時間遠隔監視システムを開発した。この章ではテレプレゼンスと遠隔監視に関する従来手法とその問題点をそれぞれ述べ、本研究の方針と概要について述べる。

## 1.1 テレプレゼンスと遠隔監視に関する従来研究

テレプレゼンスシステムとしてまず考えられる方法は、CCD カメラなどの画像センサの使用が一般的であるが、観測視野がレンズの画角により制限されるため、広範囲にわたり利用者が遠隔地を把握することは困難である。この問題を解決するために、カメラを方向が制御できる回転台の上に置いて、カメラの向きを使用者が制御しその画像を獲得する方法が考えられる (図 1)。しかし、この方法では、利用者が指示を出してからビデオカメラを実際に動かすため、機械的な時間遅延が生じてしまう。また、ビデオカメラと利用者が遠く離れている場合は、両者間の通信時間の遅延も無視できなくなる。そのほかの問題点として、多人数でテレプレゼンスを行うときに、一台のカメラでは複数人がそれぞれ異なる方向を見渡すことができなく、複数のシステムが必要となることが挙げられる。

広範囲の任意方向を注視できる手法として、静止画を利用した QuickTime VR[2] が知られている (図 2)。この手法では、あらかじめ計算機内にパノラマ画像を持っておき、その中から利用者の注視したい方向の平面透視投影画像を生成して提示する。この中で、利用者は上下・左右の任意方向を見渡せる。しかし QuickTime VR では、パノラマ画像を回転カメラや何枚ものスチル写真のつなぎ合わせで作成しており、動画像を扱うのは難しい。

時間遅れ少なく任意方向の画像が得られる例として Virtual Dome[3] が挙げられる。このシステムはカメラを常時回転させて全周の画像を生成・更新し、その画像から見ている方向の画像を切り出して利用者の装着する HMD に提示する (図 3) ため、利用者は任意方向の動画像を獲得することができる。しかし、カメラを回転させるのに時間がかかり画像が更新されるまでに時間遅延が生じる。

以上のような問題に対して、全方位画像センサ [4] を利用したテレプレゼンスが提案されている [5]。これは全方位画像センサにより遠隔地の周囲 360 度の全方位画像を獲得・伝送し、その画像より利用者に視線方向の画像を提示する。この方式では、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、広範囲の遠隔画像を見渡すことができる (図 4)。しかし、従来のシステム構成では、全方位画像センサで撮影した全方位画像を直接、計算機に伝送しており、ネットワークを介した実装には到っていなかった。

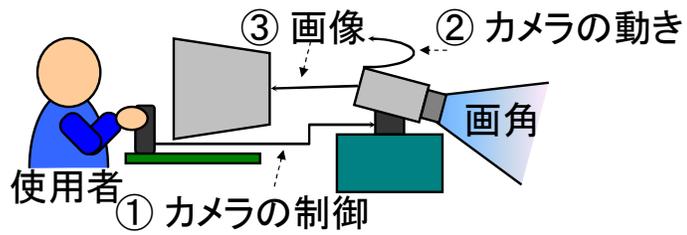


図 1 カメラの向きを制御するテレプレゼンスシステム

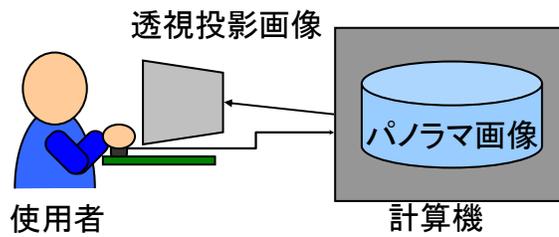


図 2 QuickTime VR

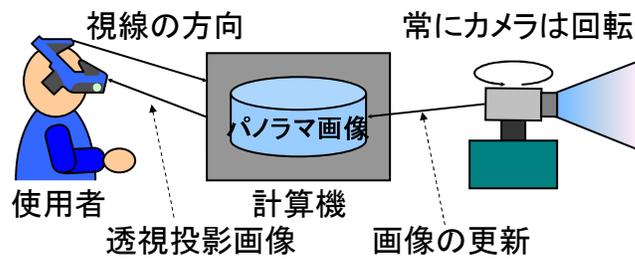


図 3 Virtual Dome



図 4 全方位画像センサを用いたテレプレゼンスシステム

従来の監視システムなどにおいては、テレプレゼンスと同様に CCD カメラが一般的に使用されている。しかし、従来の CCD カメラには、標準レンズが取り付けられたものが多く、画角が狭く撮像視野が制限されるため、環境全体を監視できない。そのため環境内を広範囲にわたって動く複数の物体の位置を全て知ることができない。この問題の解決策として、利用者が可動カメラの方向を手動でもしくは自動的に制御し、そのカメラからの画像を得る方法がある。しかし、監視者が物体の動きに合わせて手動で回転させる場合、常に画像を注視している必要があり監視者の負担になる。そのため監視者の疲労や不注意により注目していた物体を見失ってしまう可能性がある。

一方、物体の動く方向を時系列差分やオプティカルフローなどの画像処理により予測し自動的にカメラを回転させる場合、一つの物体を監視することができる [6, 7, 8]。しかし、一台のカメラでは環境全体の画像を一度に取得することはできないため、現在写っている物体の動く方向しか予測できず、複数の物体を同時に監視することができない。可動カメラを複数台使うことによって複数の物体を同時に監視することができるが、最低でも物体数以上のカメラが必要となり、監視する物体が増えるにしたがって連携制御が困難となる。また物体の位置推定を行うには物体が 2 台以上のカメラに写っている必要があり、さらに多くのカメラが必要となる。

別の解決策として固定カメラを複数使う方法がある [9, 10, 11, 12, 13]。しかしこの方法でも複数の回転カメラを使う場合と同様、カメラ間の連携制御が難しいという問題が存在する。加えて、自由に移動する物体の位置推定をロバストに行うためには死角を減らす必要があり、複数の回転カメラを使う方法と比べ多くのカメラが必要となる。また同一の移動物体の追跡もカメラを切り替えながら行う必要があり、さらに連携制御が困難となる。

もう一つの解決策として通常のカメラではなく、全方位画像センサを使用する方法が考えられる [5]。全方位画像センサを使うことにより、通常のカメラに比べ少ない台数で環境全体を監視することができる。また回転カメラと違い、カメラの可動制御を行う必要がなく扱いやすい。これまでに複数の全方位画像センサを使用して、物体の位置を推定する手法が提案されている。十河らは移動物体を人物と限



図 5 全方位画像センサと回転カメラを組み合わせた監視システム [19]

定し、その位置を一定の大きさの円として推定する手法を提案している [14]. しかし円の大きさは経験的に求めた値であり、様々な大きさの物体の位置を推定するのは困難である. さらに円の中心は三角測量により求めるので、物体がベースライン (センサ間の位置を結んだ直線) 付近に存在する場合、物体の方位角の導出誤差による推定位置の誤差が大きく、推定結果が不安定になる. さらに物体の方位角とベースライン方向が完全に一致すると、三角測量を行うことができない. また、障害物などによる部分的なオクルージョンや物体の向きなどによる物体の方位角の導出ずれにより推定位置がずれるという問題もある. このような問題に対して、寺沢らは、背景差分により移動物体を検出し、2 次元的存在領域として物体の位置推定を行う手法を提案している [15]. この手法では、物体の存在領域を背景差分領域から見える物体の方位角の範囲を使って求めるので、センサで抽出される物体の角度情報の導出誤差をそのまま反映し、様々な大きさの物体の位置推定を行うことが可能である. しかし、これらの手法では、1 台の計算機ですべてのセンサを管理しているものがほとんどで、ビル全体など多くのセンサを必要とする環境には適用できなかった.

また、全方位画像が幾何学的に歪んでいるため、図 5 のように全方位画像センサ

と回転カメラを両方使う方法も提案されている [16, 17, 18, 19]. この方法では全方位画像センサで物体の方位角を導出し, 回転カメラをその方向に向けて監視することができる. そのため全方位画像センサのみを使用した方法に比べ, 解像度の高い画像を監視者に提示することができる. しかし全方位画像センサと回転カメラという異なるカメラ間での連携が必要になる. さらに一台の全方位画像センサだけでは位置推定を行うことができない. また一台のセンサのみで 360 °あらゆる方向の画像を提示できるという全方位画像センサの特徴がなく, 複数の物体の追跡画像を提示するには物体数以上の回転カメラが必要になる.

## 1.2 本研究の方針と概要

前節の概観から, テレプレゼンスや遠隔監視においては, 以下のような要求が挙げられる.

1. 遠隔地の様子を広範囲に, かつ臨場感豊かに伝送する
  - (a) 実時間での見回し, または追跡を行う
  - (b) ユーザが遠隔地の広い範囲を見渡せる
2. ネットワークを介した通信を行う
3. 複数の物体の同時検出及び追跡を行う (遠隔監視のみ)

本研究では, 以上の要求を全て満たすネットワークを介したテレプレゼンスシステムと遠隔監視システムの開発を目的とする [20, 21, 22, 23, 24]. ネットワークを介したテレプレゼンスにおいては, ビデオレートでの全方位画像の取得により広範囲の遠隔地画像の獲得を行う. そしてネットワークを利用してその全方位動画像の伝送を行い, 実時間での観測者の視線に追従した透視投影画像の生成を行う. 本手法では観測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, ネットワークを利用して伝送された全方位動画像から透視投影画像を生成するため, 遠隔地と観測者間の距離にその時間遅延が依存することがない. したがってネットワークを介して広範囲の遠隔画像を実時間で見渡すことができる. またひとつの全方位画像センサだけで複数人が同時に異なる方向を見渡すことができ

るという特徴をもっている。

また、ネットワークを介したテレプレゼンス技術を応用した実時間遠隔監視システムでは、監視環境側をサーバ側、監視者側をクライアント側としたサーバ/クライアントモデルに基づくネットワークを利用した分散処理を行うことで、実時間監視の実現を図る。クライアント側においては、安定した複数物体の位置推定を行うために、3台の全方位画像センサを用いて監視環境(サーバ側)を撮影する。また、各サーバにおいて背景差分を利用した背景更新を行い、屋内における照明条件の変化に対してもロバストに物体検出を行うことを目指す。以上のようにして構築する遠隔監視システムはセンサ数と監視者側の数に関して、容易に拡張が可能であり、ビル全体など多くのセンサを必要とする環境に適用が可能であり、ネットワークを介した実時間遠隔監視を行うことができる。

以下、2章では、ネットワークを介したテレプレゼンスシステムについて述べ、本研究で構築した2つのプロトタイプシステムを用いて行ったそれぞれの実験結果と考察を述べる。3章では、テレプレゼンス技術の応用である移動物体注視型遠隔地画像提示システムについて述べ、そのシステムを用いて行った実験結果と考察について述べる。最後に4章で本論文のまとめとして、結論と今後の課題について述べる。

## 2. ネットワークテレプレゼンスシステム

提案するテレプレゼンスシステムは主に以下に示す三つの処理で構成される。

1. ビデオレートでの全方位画像の取得.
2. インターネットを介した全方位画像の伝送.
3. 視線追従型透視投影画像の実時間生成.

提案するシステムは peer to peer 型のシステムであり, 全方位画像センサを設置した側を送信側, 観測者側を受信側とする. 送信側で取得された全方位画像をインターネット (IPv4 または IPv6) を介して受信側に伝送する. 受信側では, 送信側から得た全方位画像より, 複数の観測者がそれぞれの観測者の視線に追従した透視投影画像を見ることができる.

本章ではまず 2.1 節において本研究で使した全方位画像センサ HyperOmni Vision について述べ, 2.2 ~ 2.4 節において上述した三つの処理についてそれぞれ述べる. 続いて, 2.5 節において提案するテレプレゼンスシステムのプロトタイプを構築し, それらを用いた実験, および考察について述べる.

### 2.1 全方位画像センサ HyperOmni Vision

図 6 に HyperOmni Vision[4] の外観とその光学系の構成を示す. 図 6(a) に示すように, HyperOmni Vision は鉛直下向きに設置した双曲面ミラー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されるものである. その双曲面には回転二葉双曲面を用いており, 図 6(b) のように 2 焦点  $O_M, O_C$  がある. 双曲面ミラーの外側にある焦点  $O_C$  にレンズ中心がくるように, カメラを鉛直上向きに取り付ける. ここで図 6(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を考えると, 双曲面ミラーの面, 双曲面ミラーの焦点  $O_M$ , ならびにカメラのレンズ中心  $O_C$  は式 (1) で表せる.

$$\begin{aligned}
\text{ミラー面} & : \frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 (Z > 0) \\
\text{ミラーの焦点 } O_M & : (0, 0, +c) (c = \sqrt{a^2 + b^2}) \\
\text{カメラレンズ中心 } O_C & : (0, 0, -c)
\end{aligned} \tag{1}$$

内側の焦点  $O_M$  に集まる像は双曲面ミラーに反射して、外側の焦点  $O_C$  に集まる。ここで、 $O_C$  にレンズ中心を置いたカメラで全方位の画像を撮影すると、3次元環境中の任意の点  $P(X, Y, Z)$  に対する全方位画像上での写像点を  $p(x, y)$  とした場合、点  $P(X, Y, Z)$  の方位角  $\theta$  は式 (2) で表される。

$$\tan \theta = Y/X = y/x \tag{2}$$

即ち  $Y/X$  で定まる点  $P$  の方位角  $\theta$  は、 $y/x$  で定まる写像点  $p$  の方位角  $\theta$  を算出することで得られる。この様に 360 度パノラマ上の領域内にある対象物体の方位角  $\theta$  が、その物体の画像面上の写像の方位として直接現れる。また、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 (3) の関係が成立する。

$$\begin{aligned}
x &= X \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \\
y &= Y \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}}
\end{aligned} \tag{3}$$

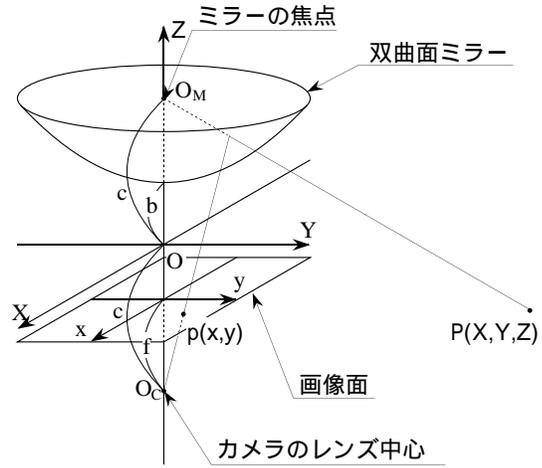
HyperOmni Vision の入力画像は双曲面ミラーの内側の焦点  $O_M$  を投影中心とした双曲面への一点中心透視投影画像であるため、入力画像 (図 7(a)) の一部を  $O_M$  から見た一般の平面透視投影画像 (図 7(b)) に変換できる。

## 2.2 全方位画像の取得

本研究において、送信側に設置された HyperOmni Vision を用いて、全方位画像をビデオレートで取得する。はじめに、ロビーや会議室のような観測したい環境全体が取得する全方位画像に含まれるような位置と高さに HyperOmni Vision を設置する。次に、取得した全方位画像をデジタルビデオストリームに変換し、送信



(a) 外観



(b) 光学系

図 6 HyperOmni Vision



(a) 全方位画像



(b) 平面透視投影画像

図 7 全方位画像と平面透視投影画像

側計算機に入力する. ここで, 図 6(a) に示すように, デジタルビデオカメラと双曲面ミラーを組み合わせた全方位画像センサを用いれば, 直接 IEEE1394 を通してデジタルビデオストリームを計算機に入力することができる.

### 2.3 全方位画像の伝送

次に, 入力されたデジタルビデオストリームをインターネットを介して, 受信側に伝送する. 本研究では, Digital Video Transport System(DVTS)[25] を用いて全方位画像を伝送する. DVTS はインターネットを介して IEEE1394 デジタルビデオストリームを伝送できるシステムであり, 以下に示すような特徴を持つ.

1. 約 33Mbps の帯域幅を使用して, 高品質動画像 (720 × 480 pixels, 30fps) を伝送することが可能.
2. DV レコーダを用いて画像データからデジタルビデオストリームへの圧縮, およびデジタルビデオストリームから画像データへの展開を行うことが可能 (図 8 参照).
3. 動画像と共にそれに対応した音声情報も伝送することが可能.
4. ネットワークの使用状況に応じた使用帯域幅の変更が可能.
5. インターネットを介してどこにでもデジタルビデオストリームを伝送することが可能.

本研究では, 送信側において, HyperOmni Vision を構成するデジタルビデオカメラを用いて全方位動画像からデジタルビデオストリームへの圧縮, および計算機を用いたデジタルビデオストリームの伝送を行う. 受信側においては, DV レコーダなどを用いてデジタルビデオストリームから全方位動画像への展開を行うことにより, 受信側計算機における負荷を抑えることが可能となる. また, 音声情報の伝送を考える場合, デジタルビデオストリームの中にデジタルビデオにより取得される環境中の音声情報が含まれるため, 音声を伝送するための他のデバイスを必要とせずに, 安価にシステム全体を構築することができる.

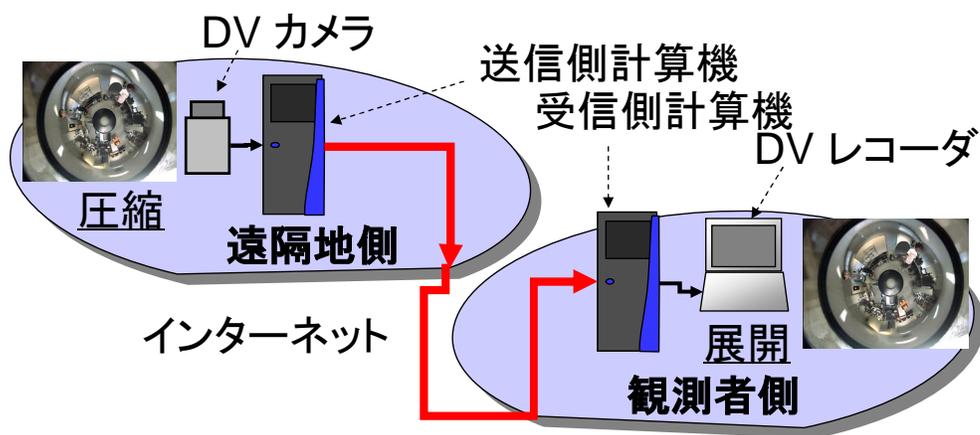


図 8 Digital Video Transport System の処理概要

## 2.4 透視投影画像の生成

受信側では, 送信側から伝送されたデジタルビデオストリームを計算機で受信し, DV レコーダなどを用いてこれを展開する. このようにして得られた全方位動画像より, 観測者の視線に追従した平面透視投影画像を生成する. ここで, テレプレゼンスにおいては視線に追従した平面透視投影画像を時間遅延なく観測者に提示することが非常に重要であるため, 2.1 節で説明したような画像の幾何学的変換を行う場合, 切り出す領域の全画素について変換計算を行うと, 計算量が膨大になりシステムの実時間性が損なわれてしまう. そこで本研究では, 格子分割による画像変形手法を採用した [5]. この手法は, 生成する透視投影画像を格子で分割し, 全方位画像上でその格子点のみをソフトウェアで正確に対応計算する. 視線に追従した実時間透視投影画像生成の概要は以下の通りである. ただし, 観測者の視点は仮想的に HyperOmni Vision のミラー焦点  $O_M$  に位置する.

1. マウスやジョイスティック, 磁気トラッカーなどのインタラクティブデバイスにより, 観測者の視線 (頭部の動き) を計測する.
2. 上記の計測に基づいて図 9 のように, 透視投影画像上に生成された格子点でのみ, 全方位画像上との対応点を求め, 切り出しを行う領域を決定する.

3. 格子の内部の画素についてハードウェアのテクスチャマッピング機能 [26] を用いた画像変形により計算を行い, 観測者の視線に追従した平面透視投影画像を生成する.

ここで, 上記の手順は連続的な観測者の視線計測にしたがって実行される. また, 複数の観測者に画像提示を行う場合, 各観測者に対してそれぞれ上記手順を実行する.

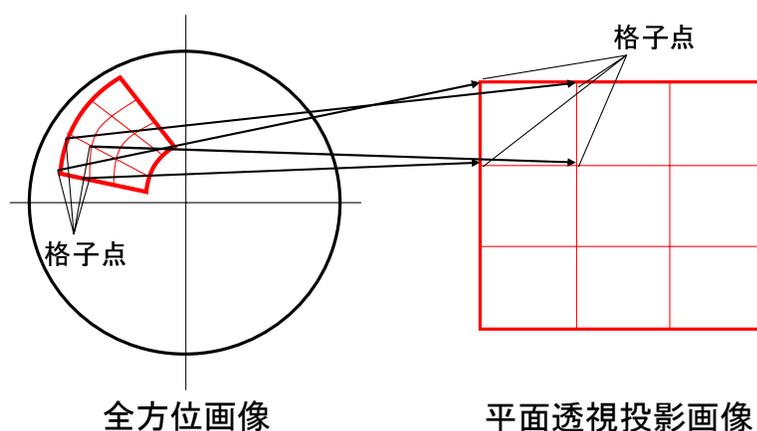


図 9 全方位画像と平面透視投影画像の関係

## 2.5 ネットワークテレプレゼンスシステムの構築

前節までに述べた技術・手法を用いてネットワークテレプレゼンスシステムのプロトタイプを構築し, 有線・無線のネットワークを用いた二つの実験を行った. 一つは学内実験として, 100Mbps Ethernet の学内 LAN のみを用いて全方位画像を伝送し, 遠隔地を複数の観測者が見回しを行った. また学外実験として, 100Mbps の学内 LAN と無線 LAN を用い一人の観測者が見回しを行った. これらの実験はテレプレゼンスシステムとしての有効性を確認するためであり, 遠隔地側と観測者側が複数のネットワークで構成された場合の動作実験である. また, 観測者の視線変化から画像提示までの時間遅延が実際の距離に依存することなく, 提案する

システムの処理構成がテレプレゼンスシステムとして有効であることを確認するためである。

以下、それぞれの実験とそれらに用いたシステムの構成、結果を述べ、それらの結果から考察を行う。

### 2.5.1 学内実験とシステム構成

本研究では、学内実験として HyperOmniVision で獲得された全方位画像を 100Mbps Ethernet の学内 LAN を用いて伝送し、受信側において送信側から伝送された全方位画像より、観測者の注視したい方向の平面透視投影画像を実時間生成し、提示するシステムを構築した。図 10 と表 1 に学内実験において構築したシステムの構成と使用した機器の仕様を示す。このテレプレゼンスシステムの送信側における DV カメラから得られる全方位画像、送信される全方位画像の解像度は共に  $720 \times 480$  pixels であり、受信側において受信される全方位画像と HMD に提示される平面透視投影画像はそれぞれ  $720 \times 480$  pixels と  $720 \times 486$  pixels (NTSC 信号) である。

図 10 に示すシステム構成で、本学情報科学研究科棟のロビー (1 階) を送信側、視覚情報メディア講座 B308 を受信側と設定した。送信側の環境中を複数の人物が自由に歩行し、図 11 のように受信側において、4 人の観測者が同時にそれぞれ異なる方向を見回す遠隔実時間モニタリングの実験を行った。

表 1 学内実験におけるネットワークテレプレゼンスシステムの機器構成

送信側・受信側計算機	CPU: Pentium4 2GHz Memory: 512MB
画像伝送システム	Digital Video Transport System <i>dvtS - 0.9a22</i>
画像表示デバイス (HMD)	OLYMPUS EyeTrek FMD-700
3D 磁気トラッカ	POLHEMUS 3SPACE FASTRAK

### 2.5.2 学内実験の結果

4人の観測者の内、2人の観測者に対するシステム実行時の提示画像と対応する入力画像を図12に示す。これらは、各観測者がそれぞれ同時に送信側環境中を歩行する2人の人物の内、別々の人物を見回しにより追跡している様子である。受信された全方位画像を図12(a)に、観測者1に提示された平面透視投影画像を図12(b)に、観測者2に提示された平面透視投影画像を図12(c)に示す。また、観測者の視線方向を3D磁気センサにより計測することによって、パン・チルトの2つの角度で生成される任意方向の平面透視投影画像を得ている。実験により、複数の観測者が音声情報とともに、送信側の固定された視点から遠隔地を自由に見回すことができた。また、遠隔地において全方位画像を取得してから、HMDへの画像提示までの時間遅延は約1秒であり、受信側においてパケットロスなしに30fpsでデジタルビデオストリームを受信することができ、提示画像も30fpsで更新することができた。また、観測者の視線変化からその方向の画像提示までの時間遅延は45msecである。ここで、前者の時間遅延は遠隔地側と観測者側との実際の距離に依存するが、後者の時間遅延は実際の距離に依存せず、本システムがテレプレゼンスを実行する上で、有効な処理構成であることが確認できる。

### 2.5.3 学外実験とシステム構成

学外実験として、学内に設置したHyperOmniVisionで獲得された全方位画像を100Mbpsの学内LANと無線LANを用いて伝送し、学外の受信側において送信側から伝送された全方位画像より、観測者の注視したい方向の平面透視投影画像を生成し、提示するシステムを構築した。図13と表2に学外実験において構築したシステムの構成と使用した機器の仕様を示す。このテレプレゼンスシステムにおいて、画像提示デバイスに4面ディスプレイを用いた。また、送信側におけるDVカメラから得られる全方位画像、送信される全方位画像、受信側において受信される全方位画像の解像度は学内実験と同様であるが、4面ディスプレイに提示される4つの平面透視投影画像はそれぞれ640×480pixelsである。

図13に示すシステム構成で、視覚情報メディア講座B315を送信側、学外の通

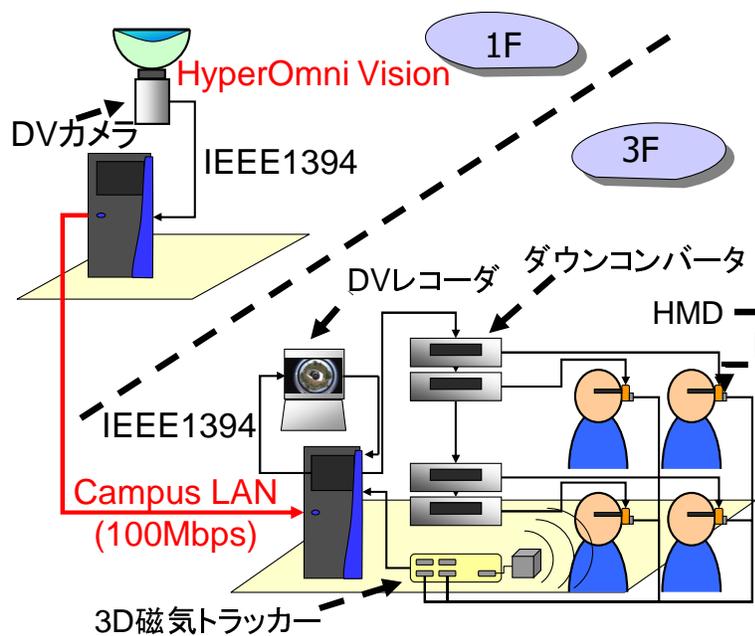


図 10 学内実験におけるネットワークテレプレゼンスシステムのシステム構成



図 11 学内実験における見回しの様子

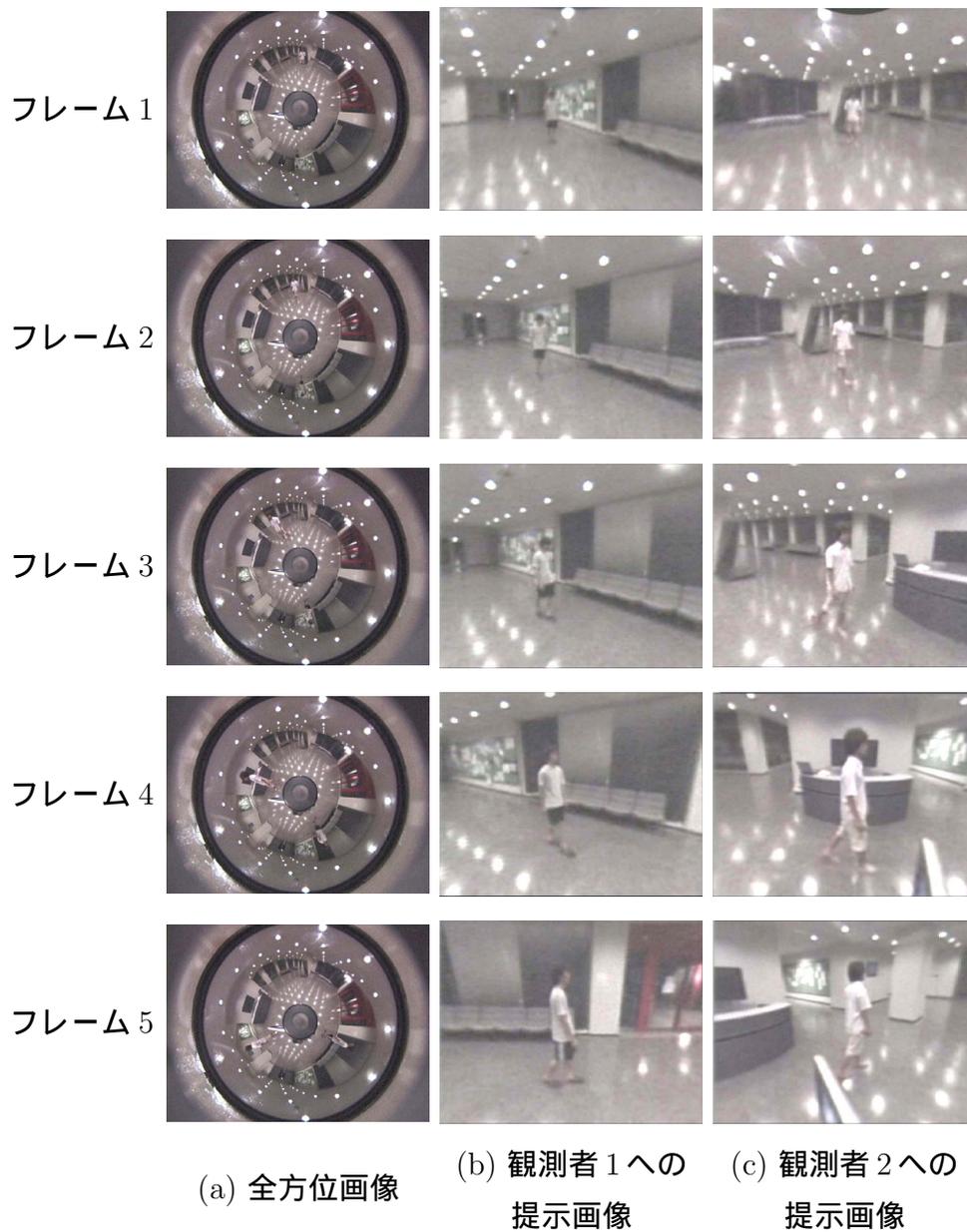


図 12 学内実験の様子

信・放送機構 奈良リサーチセンター (TAO-Nara Research Center) を受信側と設定した。送信側の環境中には複数の人物が存在し, 図 14 のように受信側において, 送信側の環境に対して 1 人の観測者がジョイスティックによる見回しを行った。

表 2 学外実験におけるネットワークテレプレゼンスシステムの機器構成

送信側・受信側計算機	CPU: Pentium4 2GHz Memory: 512MB
画像伝送システム	Digital Video Transport System <i>dvt</i> s - 0.9a22
画像表示デバイス	4 面ディスプレイ (CAVE)

#### 2.5.4 学外実験の結果

観測者に対するシステム実行時の提示画像と対応する入力画像を図 15 に示す。これは, 無線 LAN を含めたネットワークを使用した場合の実験結果であり, 遠隔地側と観測者側の距離が遠い場合を考慮したものである。受信された全方位画像を図 15(a) に, 4 面ディスプレイの左側スクリーンに提示された平面透視投影画像を図 15(b) に, 前方スクリーンに提示された平面透視投影画像を図 15(c) に, 右側スクリーンに提示された平面透視投影画像を図 15(d) に示す。実験により, 遠隔地において全方位画像を取得してから, 4 面ディスプレイへの画像提示までの時間遅延は約 1 秒であり, 受信側においてパケットロスなしに 30fps でデジタルビデオストリームを受信することができ, 提示画像も 30fps で更新することができた。また, ほとんど時間遅れなくジョイスティックの操作に対応した平面透視投影が提示され, 観測者の視線変化からその方向の画像提示までの時間遅延は生じない。

本実験より本テレプレゼンスシステムがインターネットを介した遠距離間での実時間テレプレゼンスを実現することが可能であることが確認できる。さらに, 用途に応じて様々な画像提示デバイスを選択することで, 仮想観光や遠隔通信会議, 遠隔授業支援などの実用的システムに拡張・構築が可能であると考えられる。

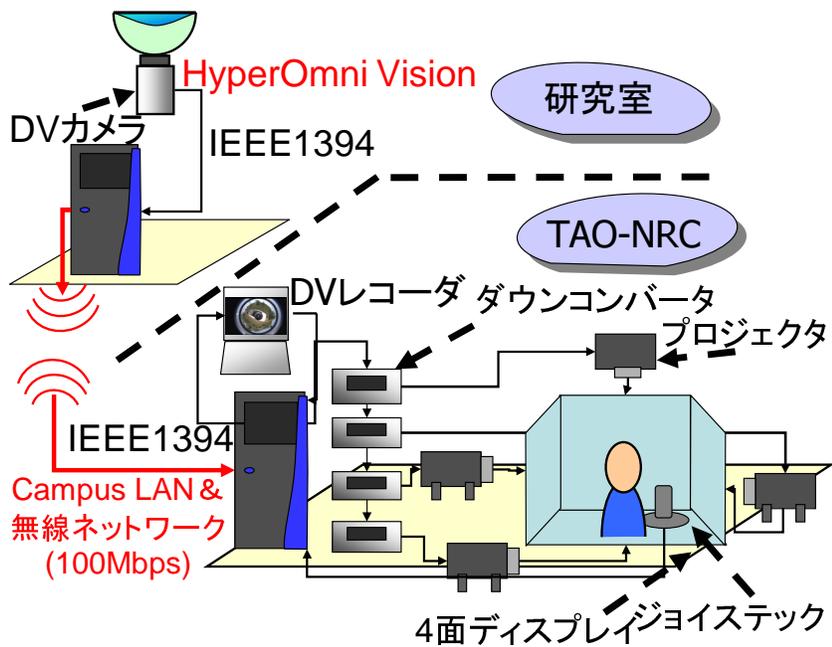
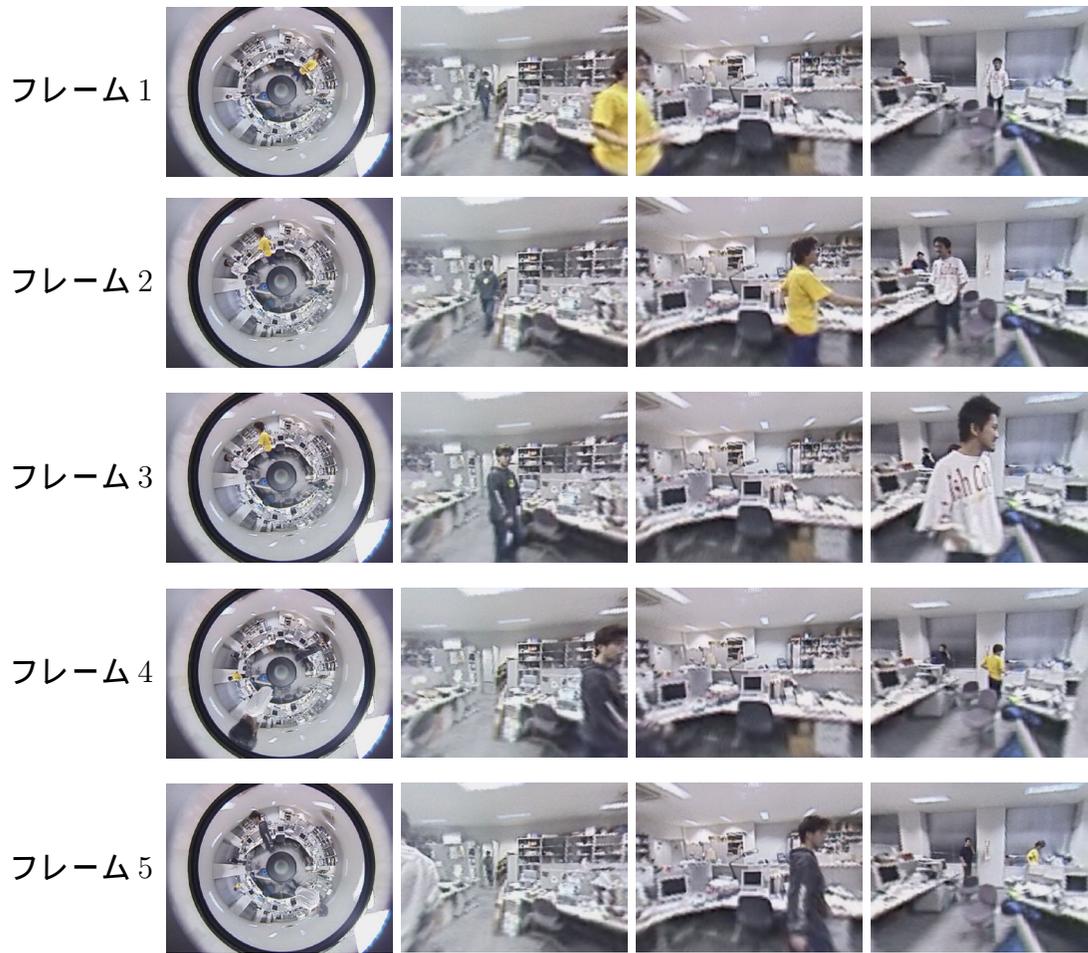


図 13 学外実験におけるネットワークテレプレゼンスシステムのシステム構成



図 14 学外実験における見回しの様子



(a) 全方位画像    (b) 左側スクリーンへの提示画像    (c) 前方スクリーンへの提示画像    (d) 右側スクリーンへの提示画像

図 15 学外実験の様子

### 2.5.5 考察

本章では、インターネットを介したテレプレゼンスシステムを提案し、それに基づいて構築した2つのプロトタイプシステムについて述べた。それらを用いて行った実験において、インターネットを介して遠隔地の全方位画像を獲得することができ、観測者は見回しに関する時間遅延が少なく、任意方向の遠隔地画像を獲得できた。また1枚の全方位画像から複数の観測者の平面透視投影画像を生成することができ、本システムがテレプレゼンスシステムとして有効であることが確認できた。しかしながら、さらに有効なテレプレゼンスシステムを構築するためには、以下に挙げられるような問題点を解決する必要がある。

1. HMD に提示される画像の鮮明度が低く、特に全方位画像センサの双曲面の中心付近に投影される領域で鮮明度が低い。
2. 30fps で全方位画像を伝送するための必要な帯域幅が大きく、ADSL や 10Mbps Ethernet, 11Mbps 無線 LAN のような一般的なネットワーク環境では、使用する帯域幅を削減しなければならない。
3. 受信側が一箇所に限定されている。

1. の問題に関しては、HyperOmni Vision の本質的な特性であり、異なった光学特性に基づいたより高精細な全方位画像センサを使用することで解決ができる。また、2. の問題に関しては、使用する帯域幅を削減し、全方位画像の伝送フレームレートを抑えると、使用者に違和感を与えてしまう。したがって、Windows Media や MPEG のような、さらに圧縮率の高い伝送方法を用いることによりこの問題を解決する必要がある。また、3. の問題に関しては、ビデオストリームのマルチキャスト機能を実現することにより解決することができる。

### 3. 遠隔監視システム

本章では、前章において提案したネットワークテレプレゼンスシステムを応用した遠隔監視システムについて述べる。ビルのような複数の部屋や複数のビルなどを一括監視する場合の遠隔監視システムを構築する条件として、以下のことが挙げられる。

1. 遠隔地に設置したカメラにより、監視環境を常に広範囲に監視する。
2. 監視環境に存在する注目すべき物体あるいはその物体の異常状態を検出する。
3. 監視環境側と監視者側との間での情報通信を行うことが可能であり、カメラの数に依存せず、実時間監視を行う。

以上の構築条件のもとで、図 16 に示すような監視システムを構築することを考える。本研究では、注目物体を、監視環境中に新たに現れた物体もしくは監視環境の中で移動する物体とする。

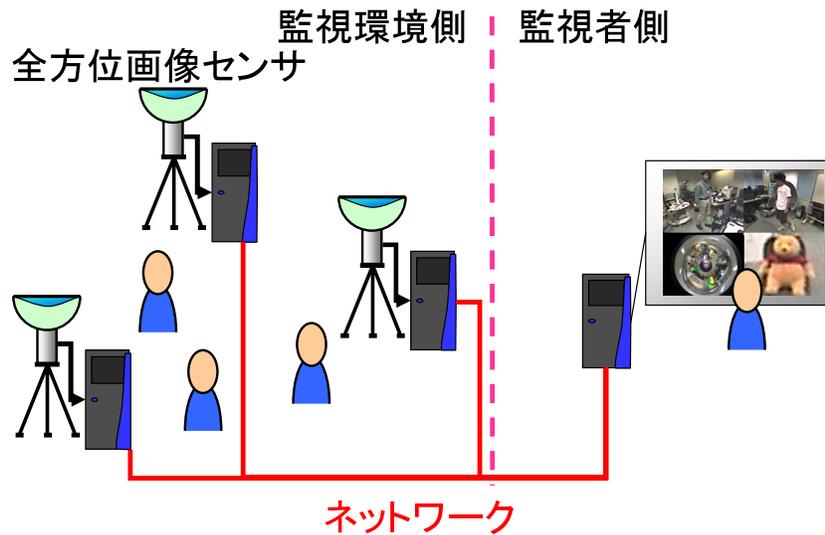


図 16 遠隔監視システムの構築図

本システムは監視環境に設置した計算機をサーバ、監視者側に設置した計算機をクライアントとしたサーバ/クライアントモデルであり、ネットワークを用いて両者間のデータ通信による分散処理を行い、ネットワークを介した実時間遠隔監視を実現する(上記構築条件 3. を満たす)。ネットワークテレプレゼンスの送信・受信側と同様に、サーバ側には全方位画像センサを 3 台と計算機を 3 台設置し、クライアント側には計算機を 1 台と DV カムコーダなどを 1 台設置し、画像提示により監視を行う。つまり、監視環境側に 3 つのサーバを設置し、常に監視環境を広範囲に監視する(上記構築条件 1. を満たす)。また、サーバはデーモンプロセスによりクライアントの要求に基づいてコネクションを確立し、ネットワークテレプレゼンスと同様に、DVTS を用いてネットワークで全方位動画像を伝送する。ただし、クライアントは 3 つの全方位画像センサのうち 1 つの全方位画像のみを受信する。各サーバにおいては、屋内における照明条件の変化を考慮した背景更新を行い、注目物体を検出する。また、クライアントはネットワークを用いて、画像データ以外の注目物体の検出情報を各サーバから受信することができる。クライアントはそれら複数のサーバから送信された検出情報を用いて、複数の注目物体の位置推定を行い(上記構築条件 2. を満たす)、監視者に物体方向の画像を提示することで追跡を行う。ここで、各サーバは常時監視環境に存在する注目物体を検出する。以下に、本システムの処理手順をサーバ側の処理とクライアント側の処理、サーバとクライアント間の情報通信処理に分けて示す。ただし、複数のサーバと複数のクライアントが存在し、1 つのクライアントは常時 3 つのサーバと接続されていると仮定する。

#### サーバ側の処理

1. 全方位画像センサを監視する環境に設置し、全方位動画像を取得する。
2. ステップ 1. で取得した全方位動画像から複数の注目物体を検出し、物体の方位情報を計算する。
3. 複数の注目物体の方位情報をクライアント側に伝送する。ただし、クライアントから画像要求のリクエストがある場合は、DVTS により画像を送信する。

ここで、上記の処理は各サーバで実行される。

## クライアント側の処理

1. 3つのサーバからの複数の注目物体の方位情報を受信する.
2. ステップ 1. で受信した各サーバからの方位情報を用いて, 監視環境における複数の注目物体の位置を推定する.
3. 推定結果に基づいて, 1つのサーバから全方位動画像を受信する.
4. ステップ 1. で受信した方位情報とステップ 3. で受信した全方位動画像より, 監視者に物体方向の画像を提示する.

## サーバとクライアント間の情報通信処理

- ・クライアント側からサーバ側への情報通信には以下のようなものが挙げられる.
  1. 全方位動画像の取得の要求.
  2. 上記 1. で受信した全方位動画像の切断の要求.
  3. 注目物体の方位情報の取得の要求.
- ・サーバ側からクライアント側への情報通信には以下のようなものが挙げられる.
  1. 全方位動画像の取得の要求に対する確認情報と動画像の送信. ただし, 動画像送信済みのサーバの場合, 動画像の送信は行わない.
  2. 全方位動画像の切断の要求に対する確認情報の送信.
  3. 注目物体の方位情報の取得の要求に対する方位情報の送信.

以上のようなサーバ/クライアントモデルにおける処理構成により, 3台のサーバを1群とした複数のサーバ群と複数のクライアントが存在する場合への拡張が容易であると考えられる.

本章では, まずサーバ側の処理として, 複数のセンサの初期設定, 監視環境下の注目物体の検出について述べる. 次に, クライアント側での処理として注目物体の存在領域推定と画像提示による追跡について述べる.

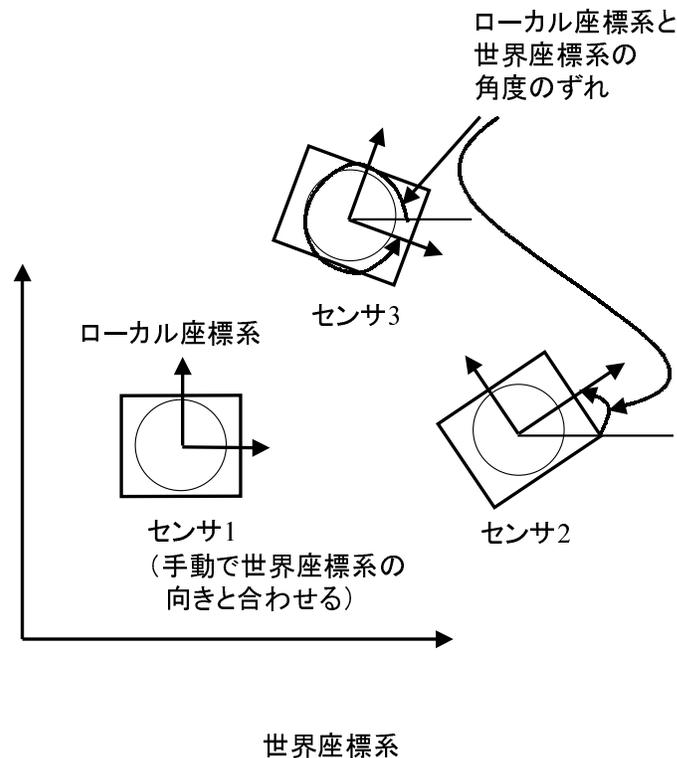


図 17 ローカル座標系と世界座標系の角度のずれ

### 3.1 センサの初期設定

センサの初期情報として、3台のセンサの環境中での位置（世界座標系における位置）を求める。ここで、世界座標とは環境のある1点（床面の角など）を原点とし各軸を床面に平行にのばした2次元座標を指す。各センサは三脚に載せ、カメラのレンズ中心とミラーの焦点を結んだ軸（図6(b)におけるZ軸）が床面と垂直になるように設置する。また実測により床面から各センサの焦点 $O_M$ （図6(b)参照）までの高さを目線あたりに一致させる。そして2つのセンサの位置を実測する。この情報と各センサから見える互いのセンサの方向から、三角測量により残り1つのセンサの位置を求める。したがって、世界座標系の高さは床面からセンサの上部まで存在する。

また、全方位画像から導出する方位角範囲の値を世界座標系での値に統一するため、各センサでのローカル座標系と世界座標系との角度のずれを求める。ここ

で、ローカル座標系とは図 6(b) の  $xy$  座標系 (全方位画像の座標系) を指す。1 つのセンサの向きをローカル座標系と世界座標系が一致するよう手動で合わせ、他の 2 センサのローカル座標系と世界座標系の角度のずれを各センサから見える互いのセンサの方向から計算する (図 17 参照)。

## 3.2 全方位画像からの注目物体の検出

監視環境内に物体が現れた場合、その物体が写った全方位画像が得られる。このとき注目物体の存在領域を推定し、物体方向の画像を提示するにあたって、全方位画像から注目物体を検出し、その方位情報を求める。以下、

- 背景更新を行う背景差分による注目物体の検出
- 注目物体の方位情報の導出

の順に述べる。ただし、方位情報とは方位角範囲と仰角範囲を意味する。また、実時間での遠隔監視、およびクライアント側での注目物体の存在領域の推定を実現するために、上記処理はサーバ側で実行され、検出した方位情報をクライアント側に送信する。

### 3.2.1 背景更新を行う背景差分による注目物体の検出

本節では注目物体をロバストに検出する手法について述べる。まず注目物体の写った全方位画像から注目物体の存在を検出する。画像から注目物体を検出する手法として、オプティカルフローによる検出 [27, 28]、テンプレートマッチングによる検出 [29]、背景差分による検出 [30] という 3 種類の手法が考えられる。

オプティカルフローによって検出を行う場合、画像から注目物体の領域のみを正確に分離するには膨大な計算量が必要となり、監視システムのような実時間性の要求される処理に適用するには問題がある。また静止している物体が検出できないという問題もある。

テンプレートマッチングによる検出手法の場合、確実に注目物体を検出するには物体の向きなどを考慮して作成した多数のテンプレートを一つ一つ入力画像に

対してマッチングする必要があり, オプティカルフローによる手法同様, 計算量が膨大になる. さらに全方位画像という円形に歪んだ画像を処理対象とするため, テンプレートの作成に工夫が必要となる.

一方, 背景差分による検出手法の場合, 画像上に現れる全ての明度変化を検出するため背景領域で明度変化の起こるような環境では誤検出が発生するが, 上記の2つの手法と比べ簡単で計算量が少なく, 実時間性の必要な処理に適した手法であると考えられる. またオプティカルフローを使った手法と違い静止している物体も検出可能であり, さらにテンプレートマッチングを使った手法と比べ検出精度に関して物体の向きなどを考慮する必要がないという長所がある.

したがって, 本研究では背景差分に基づいた注目物体の検出手法を採用した. ここでは, 蛍光灯や CRT ディスプレイなどのフリッカー, 窓の外の樹木の揺らぎなどによる明度変化やゲインアップによるごま塩ノイズなどのカメラにのみ依存する雑音による明度変化を考慮して, ロバストに注目物体を検出することを目指す [20, 21, 22].

まず入力画像中の背景画素の輝度  $I$  を以下の式でモデル化する.

$$I = \bar{I} + \sigma \sin(2\pi\omega t) + k\zeta \quad (4)$$

上式は入力画像中の各背景画素について成立し,  $\bar{I}$  は輝度値の時間平均,  $\sigma$  は輝度の振幅,  $\omega$  は輝度の周波数,  $t$  は時間,  $k$  は  $-1 \sim +1$  の値をとる係数,  $\zeta$  はカメラのみに依存した雑音を表す. ここで,  $\sigma \sin(2\pi\omega t)$  の項は蛍光灯や CRT ディスプレイなどのフリッカー, 窓の外の樹木の揺らぎなどの影響を表し,  $k\zeta$  の項はゲインアップによるごま塩ノイズなどカメラにのみ依存する雑音を表す. このとき背景画素の輝度値  $I$  は以下の範囲をとり得る.

$$\bar{I} - \sigma - \zeta \leq I \leq \bar{I} + \sigma + \zeta \quad (5)$$

本手法では, 入力画素の輝度  $I$  が式 (5) の範囲におさまる場合はその画素は背景の画素, そうでない場合は物体の画素とする.

まず, あらかじめ注目物体が存在しない状況で  $\zeta$  を測定しておく. この  $\zeta$  はカメラのみに依存する値であるので, 環境によって変更する必要はない. また, 注目物

体をロバストに検出するために、最初の  $j$  フレームにおいては  $\bar{I}$  のみを計算し、式 (5) の判定を行わない。すなわち、注目物体の検出を行わない。ただし、 $j$  は経験的に設定する値である。 $\bar{I}$  の初期値を 1 フレーム目の  $I$  とし、最初の  $j$  フレームにおいて、以下の式を用いて毎フレームごとに  $\bar{I}$  を更新する。ただし、 $\bar{I}'$  は次フレームの  $\bar{I}$  とする。

$$\bar{I}' = (n - 1)/n \times \bar{I} + 1/n \times I \quad (6)$$

n : 更新速度パラメータ

次に、背景のゆっくりとした変化を考慮し、各画素の輝度  $I$  が式 (5) の範囲におさまらず背景であると判断した場合、式 (7) を用いて毎フレームごとに  $\bar{I}$  と  $\sigma$  を更新する。ただし、 $\sigma'$  は次フレームの  $\sigma$  とし、 $\sigma$  の初期値を  $j+1$  フレーム目の  $2 \times (I - \bar{I})^2$  とする。

$$\begin{aligned} \bar{I}' &= (n - 1)/n \times \bar{I} + 1/n \times I \\ \sigma' &= (n - 1)/n \times \sigma + 1/n \times 2 \times (I - \bar{I})^2 \end{aligned} \quad (7)$$

n : 更新速度パラメータ

また、各画素が物体の部分であると判断された場合でも、急激な背景の変化や新たに監視環境中に置かれた静止物体に対応するため、式 (8) を用いて  $\sigma$  を更新する。

$$\begin{aligned} \bar{I}' &= \bar{I} \\ \sigma' &= (m - 1)/m \times \sigma + 1/m \times 2 \times (I - \bar{I})^2 \end{aligned} \quad (8)$$

m : 更新速度パラメータ ( $m \geq n$ )

これによって、新たな物体などがその場に静止し続けた場合には  $\sigma$  が増加し、しばらくするとその物体が背景と判断されるようになる。この時、 $m$  を無限大にすることで静止物体は注目物体として検出され続け、背景画像に取り込まれない。なお、本手法では、式 (5) を用いて  $I$  が背景であるか、物体であるかの判定を行うため、 $\omega$  を設定する必要はない。また、注目物体の存在しない静止環境での全方位画像を背景画像としてあらかじめ取得する必要もない。

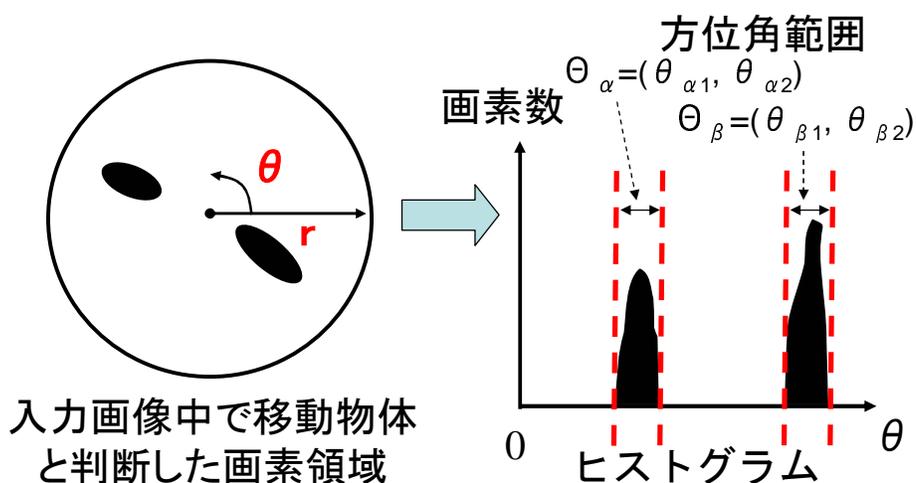


図 18 極座標変換によるヒストグラムの作成

### 3.2.2 注目物体の方位情報の導出

毎フレームごとに物体とみなされた画素領域から注目物体の方位情報を求める。本研究では文献 [15, 31] と同様に、極座標変換を利用した領域抽出により物体の画素領域を検出する。具体的には図 18 に示すように、全方位画像において直交座標から画像の中心を原点とする極座標  $(r, \theta)$  に変換し、 $\theta$  軸方向に作成したヒストグラムから連続した 0 より大きい値を持つ部分を画素領域の方位角範囲  $\theta_\alpha(\theta_{\alpha_1}, \theta_{\alpha_2}), \theta_\beta(\theta_{\beta_1}, \theta_{\beta_2})$  とする。図 18 のヒストグラムにおいて、横軸が角度  $\theta$  を、縦軸が角度  $\theta$  方向に存在する画素数を表す。極座標の  $\theta$  方向におけるヒストグラムを用いることで、背景の画素と物体の画素との領域分割における計算量が少なく、実時間性を保つことが可能であると考えられる。ただし、物体が  $0^\circ$  を含む方向に存在する場合、 $360^\circ$  の部分と  $0^\circ$  の部分をつなげた一つの連続部分を形成すると考え、方位角範囲を求める。仰角範囲に関しても同様に求める。ここで、方位情報とは全方位画像センサにおける方位角範囲と仰角範囲を表し、これらの方位情報をクライアント側に伝送する。

### 3.3 注目物体の追跡

本節では、クライアント側において、サーバ側から送られた複数の注目物体の方位情報（方位角・仰角範囲）を用い、物体の位置推定を行い、監視者に物体方向の画像を提示する手法について述べる。

本研究では、注目物体の位置推定手法に寺沢らの手法を採用した [15]。この手法は 2 次元的存在領域として注目物体の位置を推定するものであり、物体の存在領域を図 19 のような物体の方位角範囲を使って、図 20 のように求めるので、様々な大きさの物体の位置推定が可能である。ここで、注目物体の存在領域とは図 20 のように各センサの位置からのばした方位角範囲が 3 つ同時に重なる多角形領域を表す。

以下、クライアント側における注目物体の追跡を注目物体の存在領域の推定と物体方向の画像提示の 2 つの過程に分けて述べる。

#### 3.3.1 多角形領域の辺導出による存在領域推定

サーバ側から送られた複数の注目物体の方位情報（方位角・仰角範囲）の内、方位角範囲  $\theta_\alpha(\theta_{\alpha_1}, \theta_{\alpha_2}), \theta_\beta(\theta_{\beta_1}, \theta_{\beta_2})$  のみを用いて注目物体の 2 次元的存在領域の推定を行う。寺沢らは複数の物体が図 21 のような特殊な位置関係で存在すると、物体の存在しない領域が誤って抽出されるとし、その領域を誤り領域としている [15]。そして、その誤り領域の辺の導出に使用した 3 センサ全ての方位角範囲を他の領域の辺導出に使用しているとして、その条件をもとに削除を行っている。しかし、実際にそのような誤り領域に注目物体が存在する場合も容易に考えることができる。

本研究では、推定した全ての存在領域を存在領域候補とし、各存在領域候補方向を示す各センサの方位角情報を基に、それらが実際の存在領域としてどれだけ信頼できるかという信頼度ごとにクラス分けを行う。

以下、注目物体を表す多角形領域の辺導出による注目物体の存在領域推定を直線の媒介変数表現による交点座標の導出と注目物体の存在領域候補の決定、信頼度に基づく存在領域候補のクラス分けの 3 つの過程に分けて述べる。

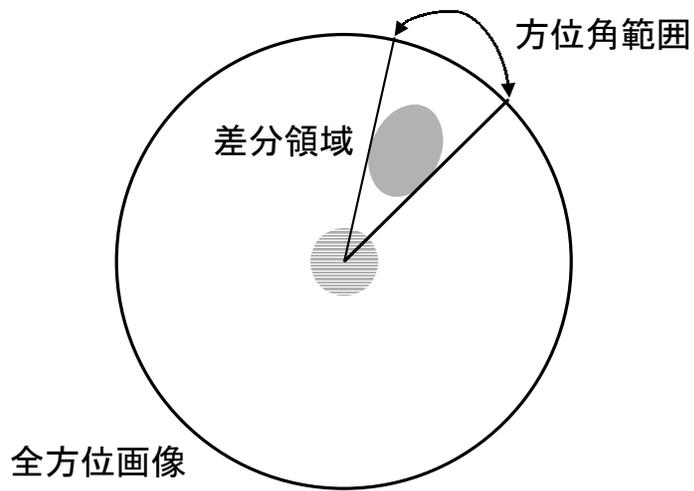


図 19 物体の方位角範囲

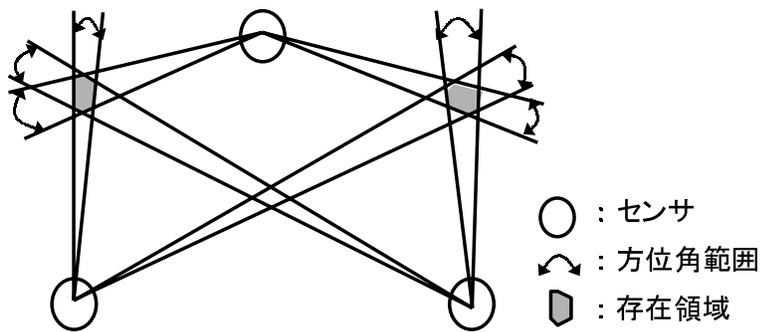


図 20 注目物体の存在領域

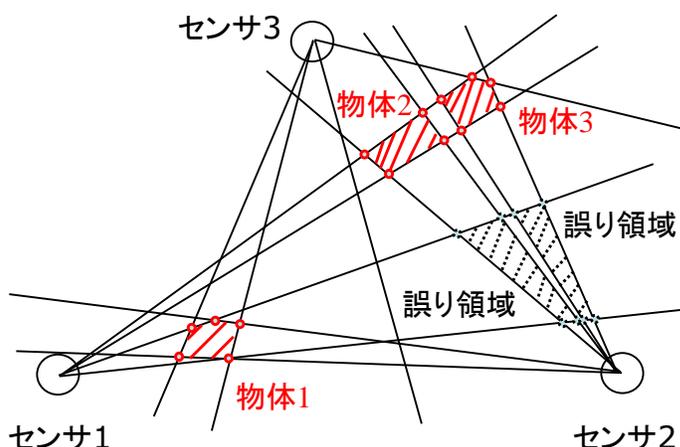


図 21 複数の人物の特殊な位置関係

### 直線の媒介変数表現による交点座標の導出

センサからのばした直線上に存在する点の座標値は、世界座標系でのセンサの位置ベクトル (センサの座標値と同値) と方位角範囲の両端の角度に対応する単位ベクトル及び媒介変数を用いて、以下の式 (9) のように求めることができる。

$$\begin{aligned}
 (\text{直線上に存在する点の座標値}) &= \\
 (\text{センサの座標値}) + (\text{単位ベクトル}) \times (\text{媒介変数}) & \quad (9)
 \end{aligned}$$

よって 2 つの直線の交点の座標値は、直線をのばしたセンサの位置ベクトルをそれぞれ  $\vec{S}_1(S_{x1}, S_{y1})$ ,  $\vec{S}_2(S_{x2}, S_{y2})$ , 直線の単位ベクトルをそれぞれ  $\vec{V}_1(V_{x1}, V_{y1})$ ,  $\vec{V}_2(V_{x2}, V_{y2})$ , 媒介変数をそれぞれ  $M_1, M_2$  とすると、以下の式で表せる (図 22 参照)。

$$\begin{aligned}
 (\text{交点の } x \text{ 座標値}) &= S_{x1} + V_{x1} \cdot M_1 \\
 &= S_{x2} + V_{x2} \cdot M_2 \\
 (\text{交点の } y \text{ 座標値}) &= S_{y1} + V_{y1} \cdot M_1 \\
 &= S_{y2} + V_{y2} \cdot M_2
 \end{aligned} \quad (10)$$

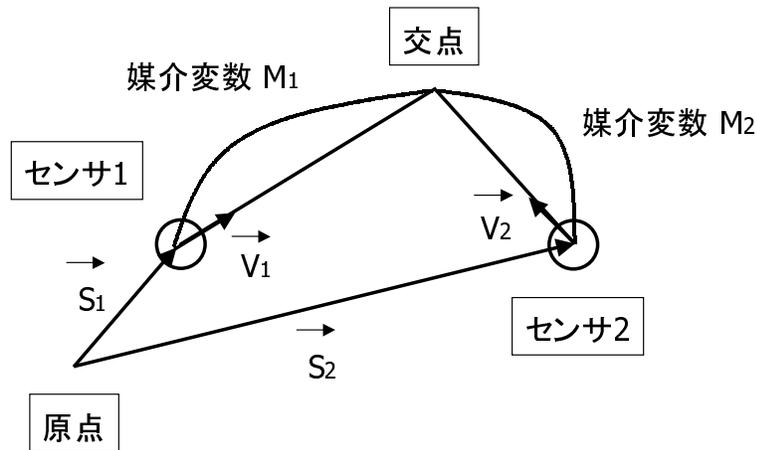


図 22 直線の媒介変数表現

以上の式より、各媒介変数は以下の式で求めることができる。

$$\begin{aligned}
 M_1 &= \frac{(S_{x2} - S_{x1}) \cdot V_{y2} - (S_{y2} - S_{y1}) \cdot V_{x2}}{V_{x1} \cdot V_{y2} - V_{y1} \cdot V_{x2}} \\
 M_2 &= \frac{(S_{x1} - S_{x2}) \cdot V_{y1} - (S_{y1} - S_{y2}) \cdot V_{x1}}{V_{x2} \cdot V_{y1} - V_{y2} \cdot V_{x1}}
 \end{aligned}
 \tag{11}$$

異なるセンサからののばした全ての 2 直線について以上の処理を行い、各交点における媒介変数を求める。

#### 注目物体の存在領域候補の決定

次に求めた交点から注目物体の存在領域候補を構成する辺を導出し、その辺を連結することにより物体が存在するであろう領域を存在領域候補として決定する。図 23 のように物体が一つだけ存在する場合、センサ 1 の方位角範囲の一端方向にのばした直線 A に注目すると、直線 A 上には他の 2 つのセンサからののばした直線との交点が計 4 つ存在する。ここで図 23(b) のように直線 A とセンサ 2, 3 から方位角範囲の両端の角度方向にのばした直線との交点によってできる辺を B, C とすると、直線 A 上に存在する存在領域を構成する辺は辺 B と C の重なる部分 (図 23(b) の太線) となる。そこで両端点の媒介変数により直線 A 上で辺 B, C の取りうる媒介変数の範囲を求め、その重なり判定により存在領域を構成する辺の両端

点の媒介変数を求める。最後に全ての直線に対して求めた辺を連結することにより、一つの閉じた領域として存在領域候補を求める。ただし、物体がベースライン方向に存在する場合やセンサの近くに存在し、その方位角範囲が広い場合など、注目する直線、交わる相手の直線に対応する媒介変数のうちどちらか一方でも負になる点は媒介変数範囲に含まれず、存在領域候補の辺にも含まれない。また、物体が複数存在する場合、物体毎に辺を連結する必要がある。連結する辺は同一点を端点として共有しているので、センサからのばした直線に対してラベル付けを行い、辺の両端点において交わる2直線のラベルをもとに辺を順次連結していくことができる。

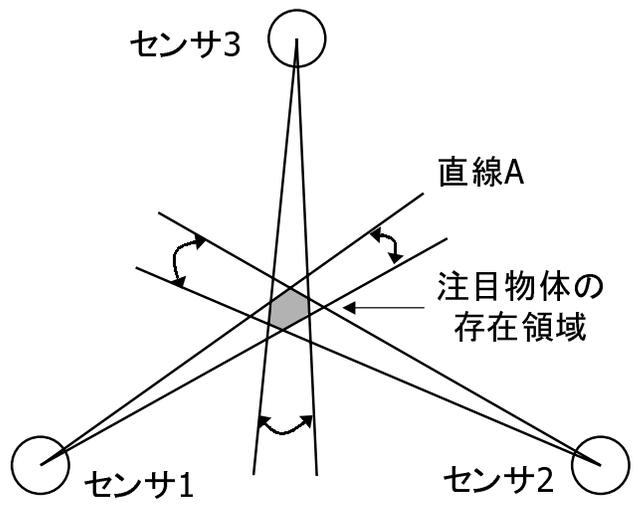
#### 存在領域候補のクラス分け

複数の物体が特殊な位置関係で存在すると、物体数以上の存在領域候補が存在し、物体の存在しない領域が誤って検出される。図21のように、3つの物体の特殊な位置関係によって誤り領域が発生し、寺沢らの手法ではこれらは先に述べた条件より削除している。しかし、実際に5つの物体がそれぞれの存在領域候補に存在した場合、誤り領域に存在する物体を検出することができない。

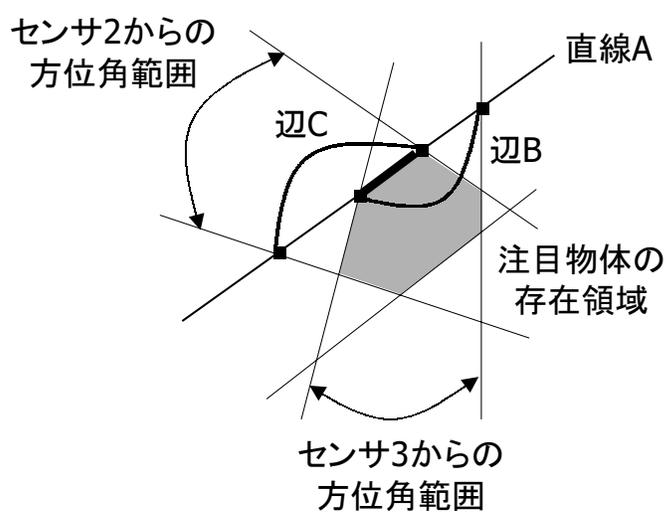
本研究では検出した注目物体の存在領域候補方向を示す各センサの方位角範囲を基に、それらが実際の存在領域としてどれだけ信頼できるかという信頼度ごとにクラス分けを行う。ここで、信頼度とは各センサの方位角範囲にある存在領域候補数の逆数であり、一つの存在領域候補は各センサからの3つの信頼度を持つと考える。例えば、図24における物体1の領域はセンサ1の方位角範囲(1)に関しては $1/3$ 、センサ2の方位角範囲(2)に関しては $1/1$ 、センサ3の方位角範囲(3)に関しては $1/1$ の信頼度を持つと考える。以下に存在領域候補の信頼度に基づく各クラスを示す。ただし、クラス1が最も物体の存在する可能性が高く、次いでクラス2、クラス3と順序立てる。

##### ・クラス1)

図25における赤色領域。存在領域候補が持つ3つの信頼度の内、信頼度が $1/1$ であるものが少なくとも一つ存在する領域。



(a) 全体図



(b) 拡大図

図 23 1つの物体の存在領域

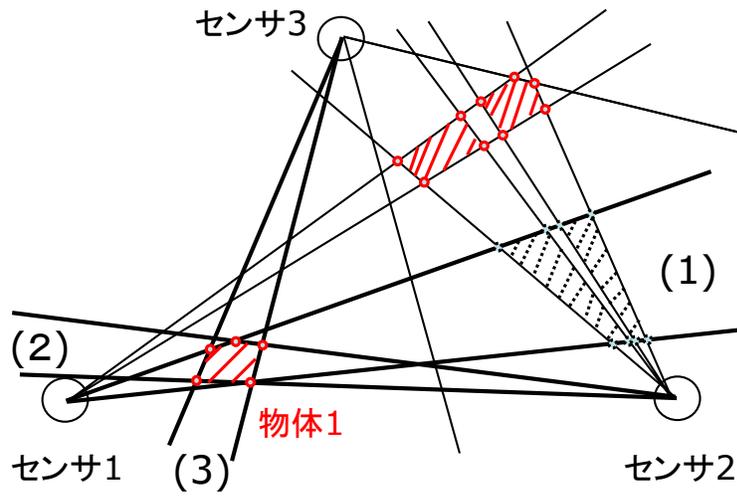


図 24 物体 1 の信頼度

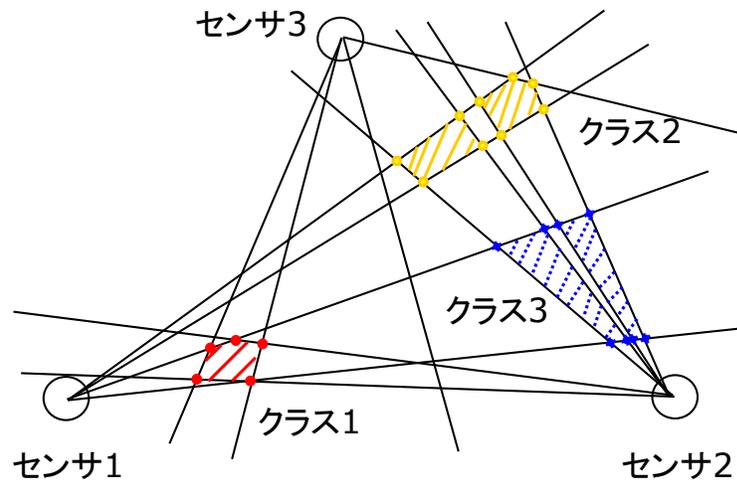


図 25 存在領域候補のクラス分け

・クラス2)

図 25 における黄色領域. 存在領域候補が持つ 3 つの信頼度全てが  $1/2$  以下であり, かつ各方位角範囲にクラス 1 の存在領域候補が存在しない場合.

・クラス3)

図 25 における青色領域. 存在領域候補が持つ 3 つの信頼度全てが  $1/2$  以下であり, かついずれかの方位角範囲にクラス 1 の存在領域候補が存在する場合.

全ての存在領域候補を信頼度に基づいて, 上記のようなクラスに分類し, それらを監視環境内に存在する注目物体の存在領域とする.

### 3.3.2 注目物体方向の画像提示

最後に, 受信する全方位動画像から複数の物体方向の平面透視投影画像を生成し, 監視者に提示する.

ここで, 全方位動画像の取得要求を送信するサーバは監視者の手動で選択されるものとする. まず, クライアントは選択したサーバに全方位動画像の取得要求を送信し, サーバから全方位動画像を受信する. 次に, 受信した全方位動画像とそれに対応した注目物体の方位情報を用いて, 物体方向の平面透視投影画像を生成する. 全方位動画像から平面透視投影画像へはネットワークテレプレゼンスと同様に参考文献 [5] の手法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変換する. ただし, 仰角範囲を基に, 図 6(b) におけるミラーの焦点  $O_M$  と変換画像間の距離 (仮想的な焦点距離) を変化させることで, 一般のカメラのズーム機能を実現する.

## 3.4 遠隔監視システムの構築

本節では, 本研究で提案した手法を用いて構築した遠隔監視システムのプロトタイプについて述べ, 次に, 注目物体の検出及び画像提示による追跡実験について述べる.

### 3.4.1 システム構成

図 26 と表 3 に遠隔監視システムの構成図と使用した機器の仕様を示す。構築した遠隔監視システムでは、本学情報科学研究科 B 棟 3 階 視覚情報メディア講座 B308 を監視環境としてサーバシステムを設置し、B211 を監視者側としてクライアントシステムを設置する。また、使用したネットワークは有線 100Mbps の学内 LAN を使用した。サーバ側では同一直線上に 3 台の全方位画像センサが並ばないようにそれらを設置する。各センサで取得した全方位画像をそれぞれに接続された計算機に入力し、サーバ側で注目物体の検出を行い、クライアント側で物体の位置推定を行い、画像提示による追跡を行うものである。ただし、全てのサーバシステムにおいて、非同期で全方位画像の入力及び注目物体の検出が実行される。

### 3.4.2 人物の検出・追跡実験

構築したシステムを用いて、3 種類の注目物体の検出及び追跡の実験を行った。以下、それぞれの実験の設定、結果について述べる。

#### 実験の設定

監視環境は  $9.2m \times 9.2m$  の大きさの実験室であり、正方形の形をしているので、床面の角の 1 頂点を原点とし、この点を含む 2 辺を  $x$  軸、 $y$  軸とする座標を世界座標とした。そして 3 台の全方位画像センサを  $(0.0m, 0.0m)$ 、 $(4.5m, 0.0m)$ 、 $(3.0m, 4.5m)$  の位置に配置した。また床面からミラー焦点までの高さを  $1.7m$  とした。実験環境を図 27 に、またそれぞれの実験の詳細を以下に示す。

#### ・実験 1 )

1 人の人物の位置推定実験。人物は設定したルートを歩行することで、正しく注目物体の存在領域が検出されていることを確認する。なお、設定したルートは世界座標において  $(1.5m, 0.0m) \rightarrow (3.0m, 0.0m) \rightarrow (3.0m, 2.0m) \rightarrow (1.5m, 2.0m) \rightarrow (1.5m, 0.0m)$  の順に直線移動するルートであり、人物は通常の人間の歩行速度に近い約  $80cm/s$  で歩行する。

表 3 遠隔監視システムの機器構成

計算機 (サーバ・クライアント)	CPU: Pentium4 2GHz Memory: 512MB
画像伝送システム	Digital Video Transport System <i>dvtS - 0.9a22</i>

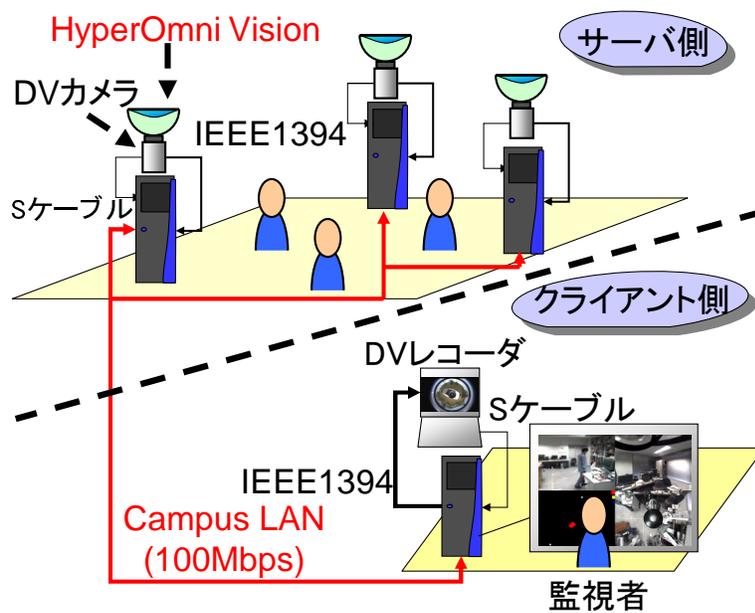


図 26 遠隔監視システムの構成

・実験 2 )

1 人の人物が監視環境を自由に歩いた状態での注目物体の検出及び追跡.

・実験 3 )

2 人の人物が監視環境を自由に歩行し, さらに 1 つの静止物体を監視環境に置いた状態での注目物体の検出及び追跡.

実験 2 と 3 では, 人物はドアから監視環境に入り, 自由に移動した. また, 実験 3 においては一人の人物が一つの静止物体を環境中に新たに置いた. なお, サーバにおける注目物体の検出において, 本実験では  $0 \leq I \leq 255$  に対して,  $\zeta$  は 4 とした. また,  $j$  は 360,  $n$  は最初の 360 フレームまではフレーム番号, 361 フレーム以降は 360 とし,  $m$  は 1080 とした. ここで  $n$  と  $m$  はいくつかの実験から決定されたヒューリスティックな値であるが, ビルのロビーや会議室などの一般的な環境に適用できると考える.

クライアント側で監視者に提示される画面の例を図 28 に示す. 本実験では画面を 4 分割し, 左下に位置推定結果を, 残り 3 つに検出された物体を入力画像中で面積の大きい方から提示した. ただし, 注目物体が 3 つ未満の場合は, あらかじめ決められた速度で自動的に環境を見回す画像や図 6(b) に示す全方位画像センサのミラー焦点から鉛直下向きの画像を提示するように実装した. なお, 位置推定結果において, 黒色の小さな正方形はセンサの位置, 画像全体が監視環境を表し, 右上の赤, 黄, 青は図 25 において述べた検出された存在領域のクラスを表しており, それぞれクラス 1, クラス 2, クラス 3 に対応する.

## 実験結果

まず実験 1 として, 1 人の人物の位置推定実験を行った. 推定した全ての存在領域を描画した CG を図 29 の多角形領域 (ワイヤーフレーム) で示す. 図 29 において, 緑色で書かれた線が人物が目標とした歩行ルートを示す. また, 位置推定実験時の各サーバの入力全方位画像と注目物体の検出結果の内, 選択したフレームを図 30 に, クライアントにおいて提示した平面透視投影画像の内, 図 30 と同一のフレームを図 31 に示す. ここで, 図 30 は, 各センサの入力全方位画像に CG で描画

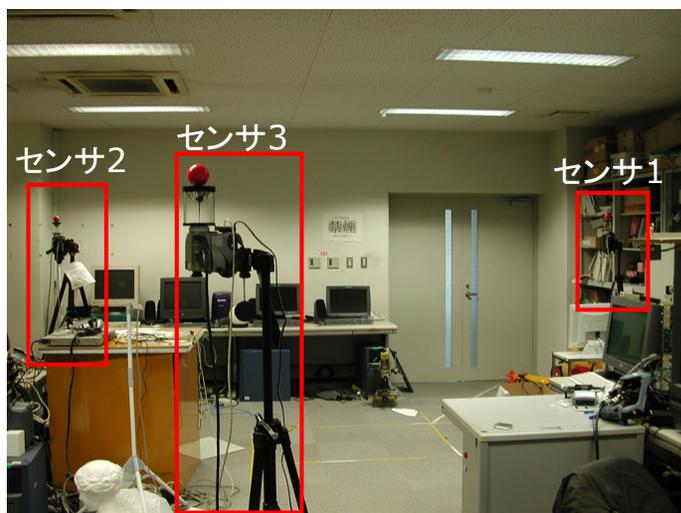


図 27 実験環境



図 28 監視者への提示画像

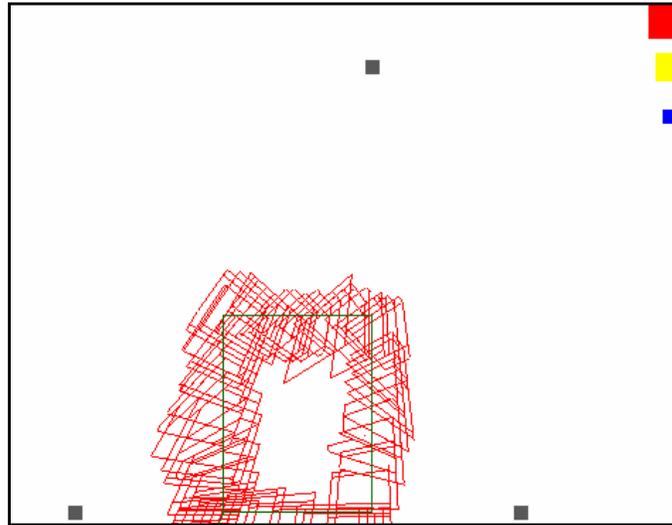


図 29 実験 1 で推定した存在領域の連続 CG

した物体の検出結果を重畳表示したものであり、赤色と緑色の直線で囲まれた領域が検出された注目物体を示す。

また実験 2 として、1 人の人物が監視環境を自由に歩いた状態での注目物体の検出及び追跡を行った。各サーバの入力全方位画像と注目物体の検出結果の内、選択したフレームを図 32 に、クライアントにおいて提示した平面透視投影画像と CG で描画した位置推定結果の内、図 32 と同一のフレームを図 33 に示す。また実験 1 とは異なり、推定した注目物体の存在領域を各クラスの色で塗りつぶされた多角形領域で示す。

次に実験 3 として、2 人の人物が監視環境を自由に歩行し、さらに 1 つの物体を環境に置いた状態での監視を試みた。各サーバの入力全方位画像と注目物体の検出結果の内、選択したフレームを図 34 に、クライアントにおいて提示した平面透視投影画像と位置推定結果の内、図 34 と同一のフレームを図 35 に示す。また全ての実験において、サーバ側における物体検出により方位情報は約 0.2 秒間隔で更新できた。またクライアント側においては人物の方位情報は約 0.2 秒ごとに更新

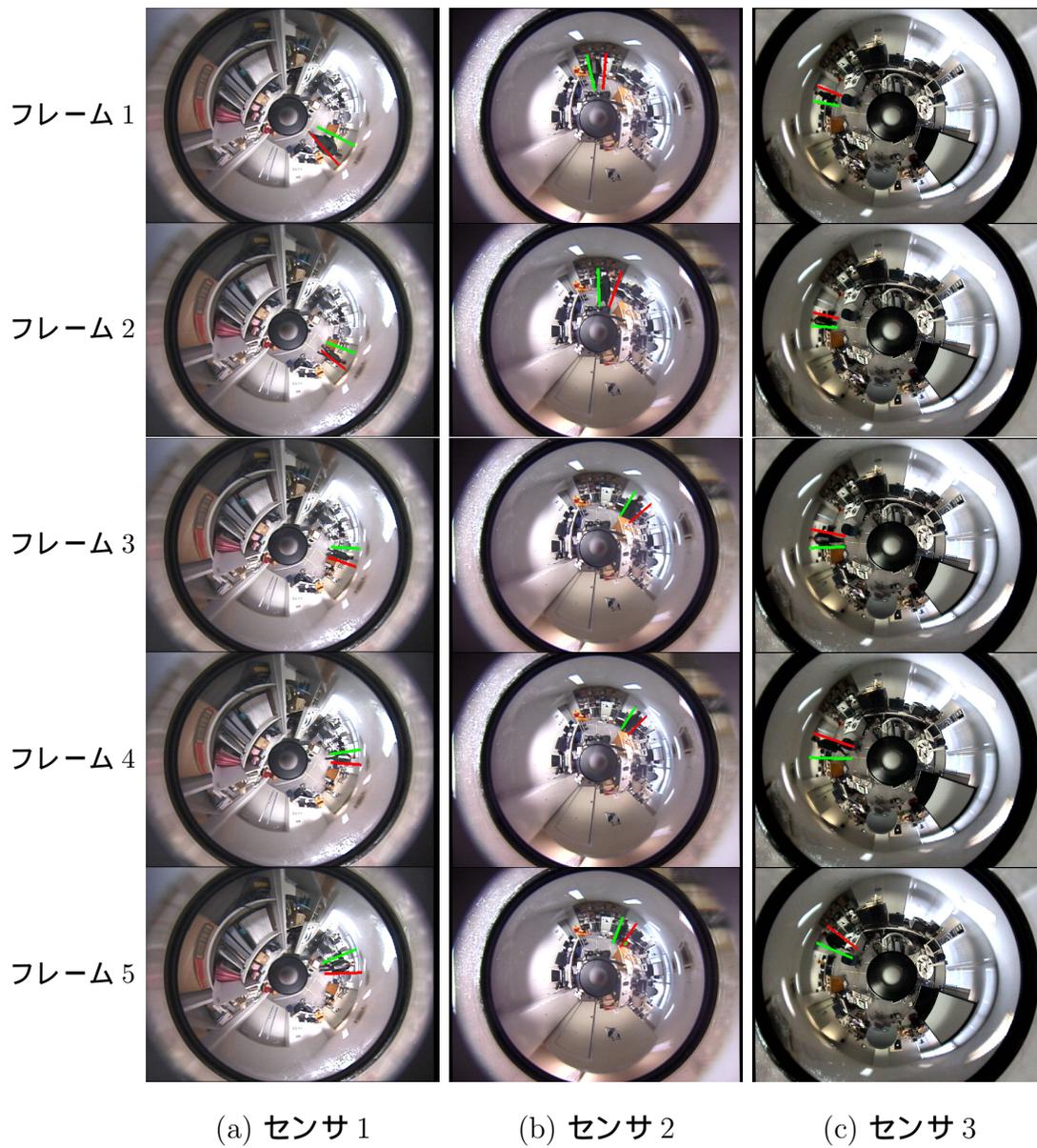


図 30 実験 1 でのサーバにおける入力画像と注目物体の検出

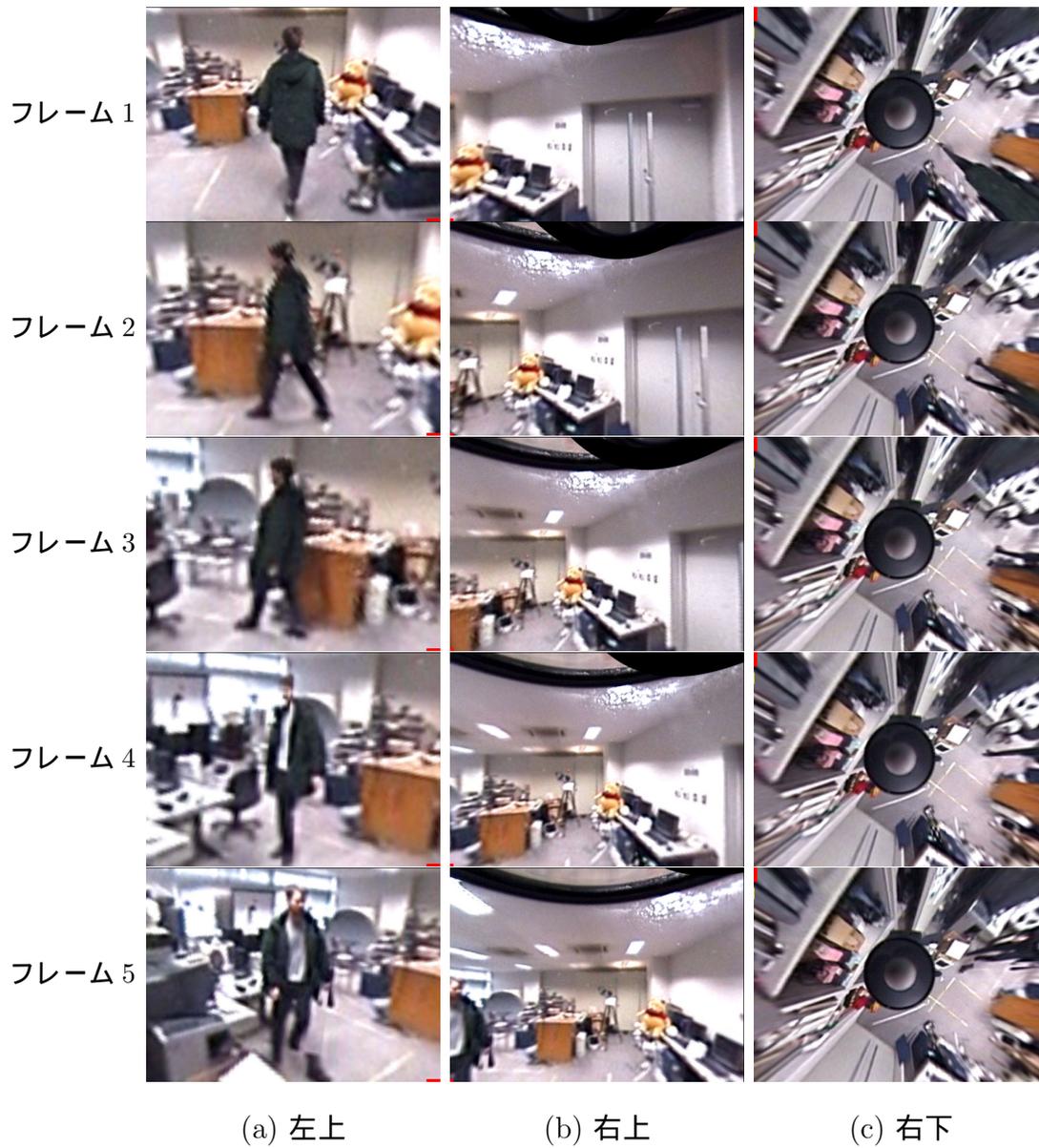


図 31 実験1でのクライアントにおける提示画像 (図 28 を分割)

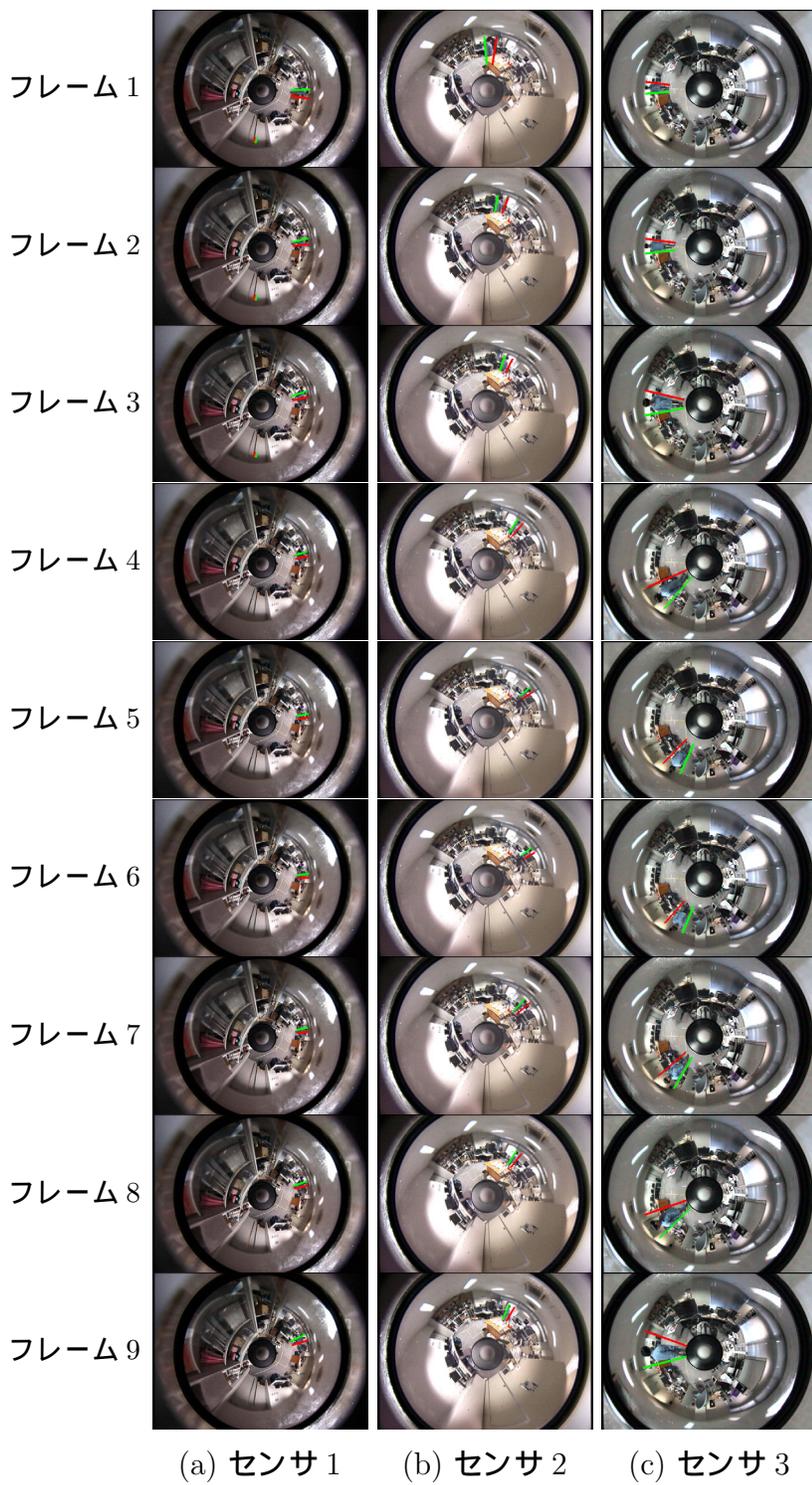


図 32 実験 2 でのサーバにおける入力画像と注目物体の検出

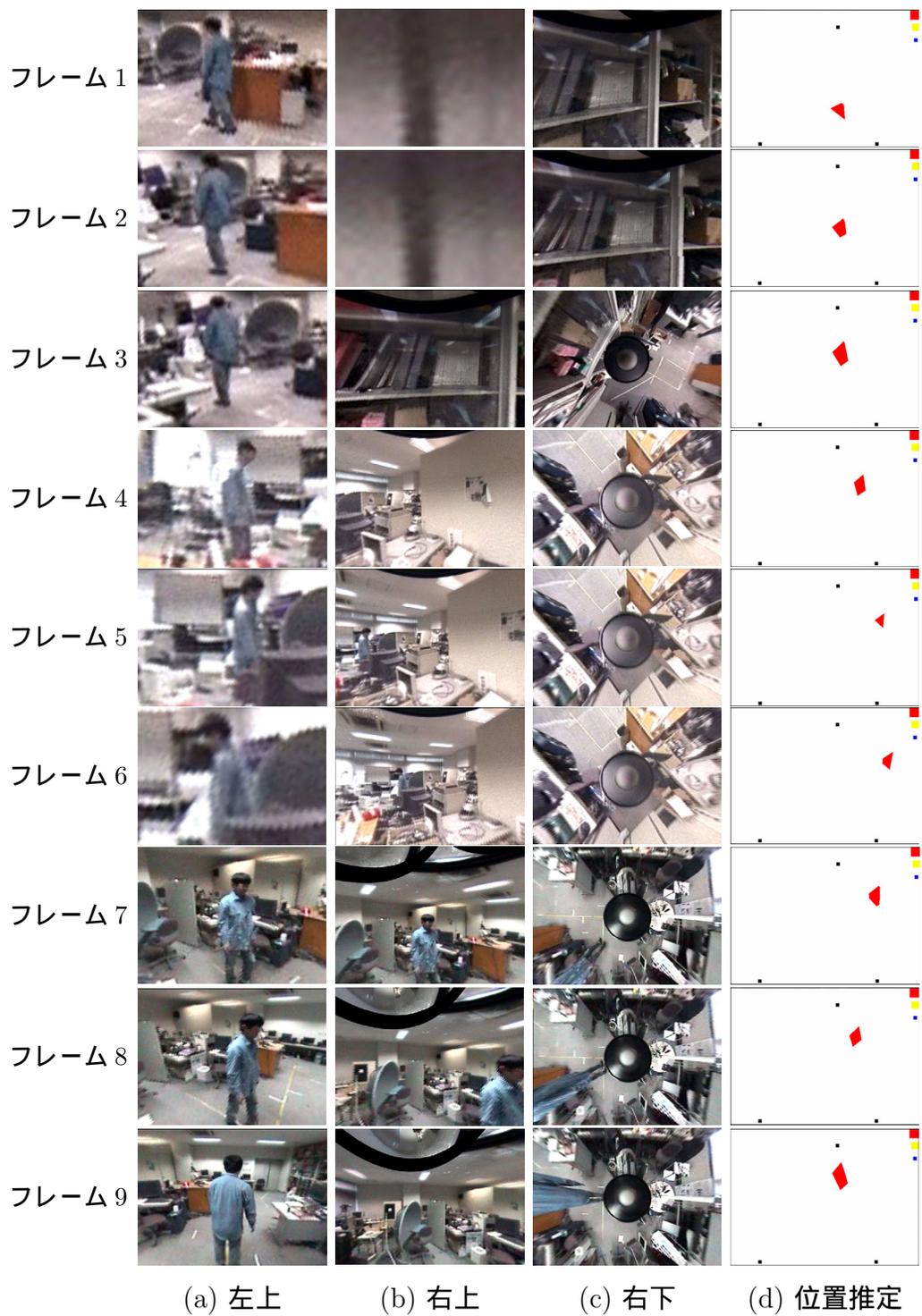
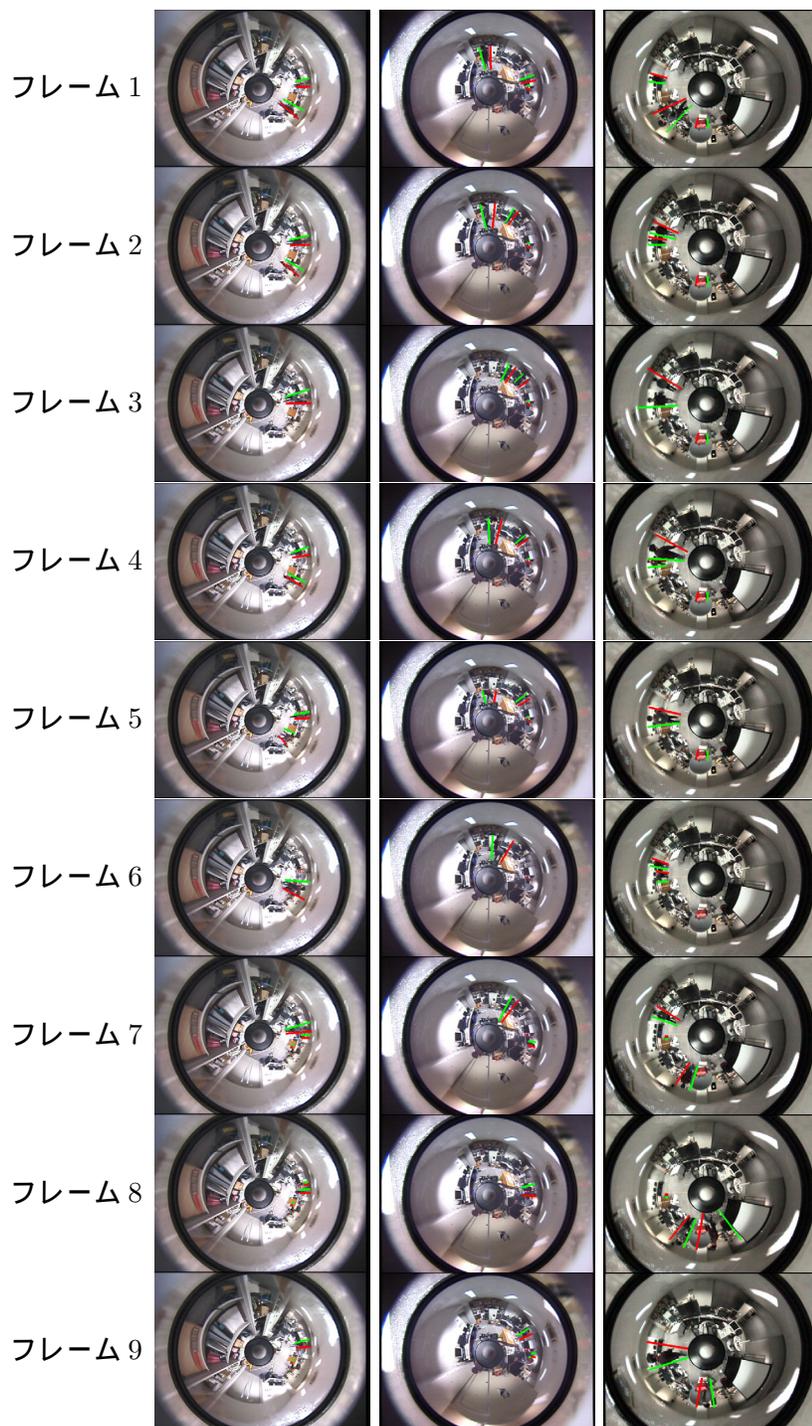


図 33 実験 2 でのクライアントにおける提示画像 (図 28 を分割)



(a) センサ 1    (b) センサ 2    (c) センサ 3

図 34 実験 3 でのサーバにおける入力画像と注目物体の検出

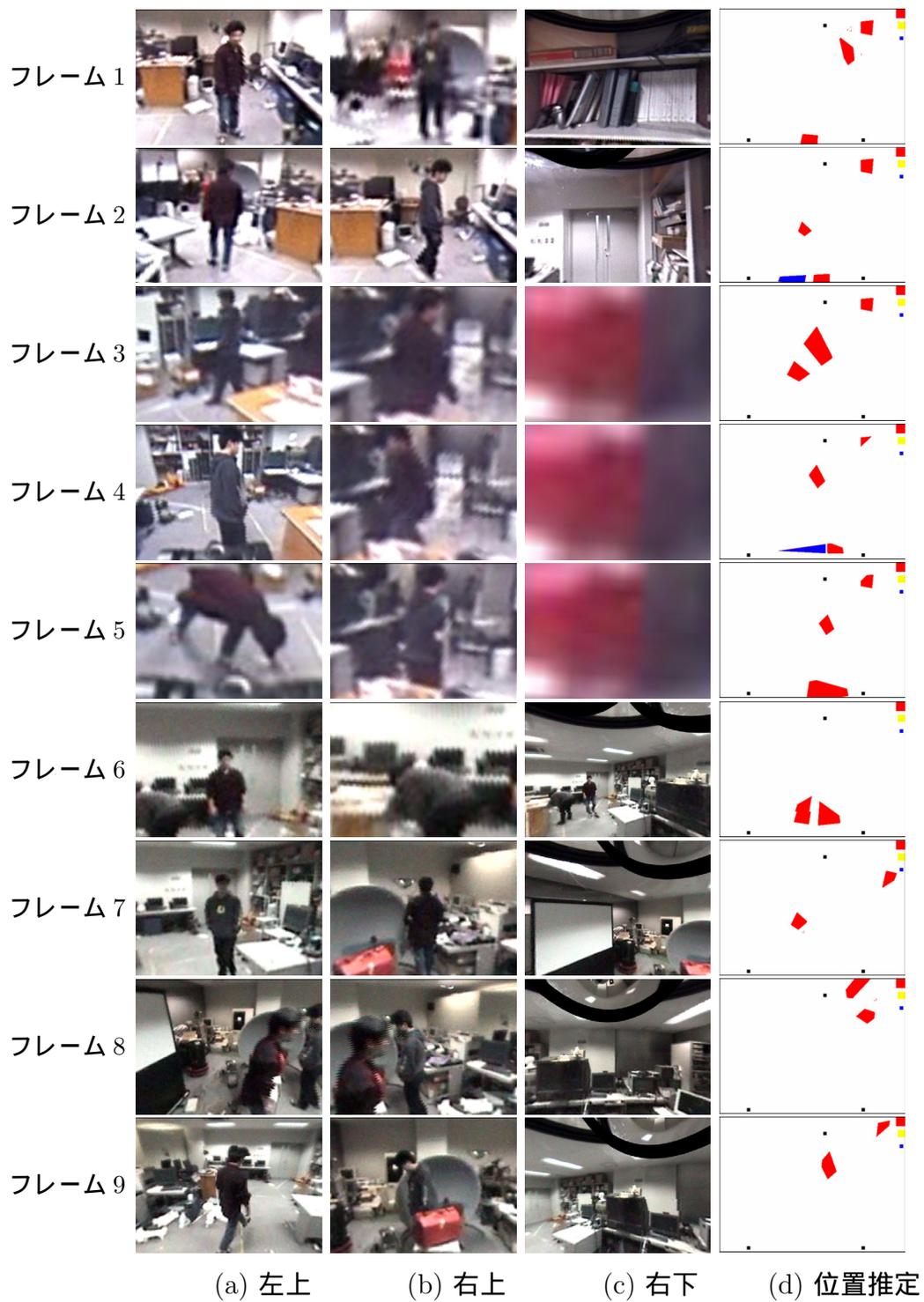


図 35 実験 3 でのクライアントにおける提示画像 (図 28 を分割)

され、それに基づいた移動物体の位置推定を行うことができた。さらに人物の方位情報は約 0.2 秒ごとに更新されるものの、提示画像は約 0.05 秒ごとに更新できた。なお、遠隔地における全方位画像の取得から画像提示までの時間遅延は約 1 秒であり、クライアント側においてパケットロスなしに 30fps でデジタルビデオストリームを受信することができた。

### 3.4.3 考察

実験 1 の結果について考察する。クライアント側では各サーバにおいて非同期に検出された物体の方位情報を用いて、複数の物体の位置推定を行う。したがって、クライアント側での処理フレームごとに各センサの方位角範囲の時間的なずれが位置推定に影響を及ぼすことが考えられる。しかし、図 29 より設定した移動ルートに沿って人物の存在領域が推定できていることが確認できる。したがって、通常の歩行速度で移動する物体に対して、その存在領域の推定が可能であると考えられる。また、図 30 より、各サーバにおいて、蛍光灯などのフリッカーは無視し、1 人の人物が検出され、物体の方位情報が検出されていることが確認できる。また、図 31 より、各サーバから送信された方位情報を基に、物体方向の透視投影画像を作成・提示できていることがわかる。ここで、監視環境に存在する注目物体が 1 つであるため、図 31(b) と (c) は物体方向の画像を提示しておらず、それぞれ一定の速度で自動的に環境を見回す画像と決められた方向の画像を提示していることが確認できる。

次に実験 2 の結果について考察する。実験 1 と同様に、図 32 より、各サーバにおいて、監視環境を自由に歩く 1 人の人物のみが検出されていることが確認できる。また、図 33 より、各サーバから送信された方位情報を基に、物体方向の透視投影画像を作成・提示できていることがわかる。しかし、図 32(a) のフレーム 1~3 や図 33(b) のフレーム 1~2 に示されるように、人物が壁や窓に近づくとその部分に影が写りこみ、その部分が抽出され、人物の方位情報と別に影のみの方位情報が検出され、結果として影の存在領域が別に推定される場合があった。また、影が非常に大きく写りこみ、存在領域が大きく抽出される場合もあった。

最後に実験 3 の結果について考察する。図 34 と図 35 において 2 人の人物と新たに置かれた物体が検出できていることがわかる。また、一定の時間存在した赤色

の物体は図 34 と図 35 のフレーム 6 以降, 背景と判断され, 注目物体として検出されていないことも確認できた. 図 35 のフレーム 1 とフレーム 2 では注目物体が 2 つ検出されている. これは図 35 のフレーム 1 における (b) の画像とフレーム 2 における (a) の画像において, 赤色の静止物体と人物がセンサ 1 から見て重なっており, 1 人と判断されているためである. また, 図 35(d) から, 各サーバから送信された複数の注目物体の方位情報を基に, 複数の物体の存在領域を同時に推定できていることがわかる. 人物が複数存在する場合, 図 35(d) のフレーム 2 と 4 の推定領域の CG からわかるように, 2 人の人物の特殊な位置関係により, 存在確率が小さい青色の領域が生じる. しかし, フレーム 2 と 4 それぞれの (a) と (b) の提示画像より, その領域に物体が存在しないことが画像提示により確認できる. また, 蛍光灯などのフリッカーなどは無視し, 複数の注目物体のみを検出できることが確認できる.

以上のように, 本研究で構築した遠隔監視システムは監視環境内に存在する複数の人物を同時に検出し, ネットワークを介して監視者が画像提示による追跡を行うことが可能である. さらにサーバ側における背景更新により, 監視環境内に一定の時間存在した静止物体などは背景として判断され, 注目物体として検出されない. 照明条件などによる背景の変化がある場合でも, ロバストに注目物体の検出を行うことができる. また, ネットワークを介してサーバ側とクライアント側に処理を分散させることで実時間での提示画像の更新を行うことができ, ビル全体などの多くのセンサを必要とする環境を監視する場合, 有効なシステムであると考えられる. しかし, 構築した遠隔監視システムはネットワークテレプレゼンスと同様に, 全方位動画像を伝送するために 33Mbps の帯域幅を使用しているため, 更なる帯域幅の削減が望まれる. また, サーバ側とクライアント側での方位情報と提示画像の更新速度が異なるために, クライアント側において方位情報が更新されない処理フレームが存在する. そのため, 提示画像においてなめらかに注目物体を追従することができない. その他, 監視環境中に複数の注目物体が存在した場合に, フレーム間での同一物体の対応付けを行っていないなどの問題点が挙げられる.

## 4. おわりに

本研究では、全方位画像センサによるネットワークを介したテレプレゼンスシステムと遠隔監視システムを開発した。構築したテレプレゼンスシステムはビデオレートでの全方位画像の取得、ネットワークを利用した全方位画像の伝送、視線に追従した透視投影画像の実時間生成から構成され、実験より、観測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、ネットワークを介して、広範囲の遠隔画像を見渡すことができた。提案システムはさまざまな分野で応用が期待できる。例えば、遠隔操作支援、遠隔観察・監視、遠隔通信会議・授業支援、仮想観光、水族館テレプレゼンスなどでの実利用が考えられる。また、テレプレゼンス技術の応用として開発した遠隔監視システムは3対1のサーバ/クライアントシステムである。さらに、ネットワークを利用して多対多のシステムへの拡張も考えたネットワークプロトコルと分散処理を行うことでさらに大規模なシステムを考えた場合にも、実時間での提示画像の更新・監視を行うことができると考える。サーバ側においては、背景差分を利用し、かつ背景の更新を行うことで、複数の注目物体を同時に検出できた。またクライアント側においては、サーバ側から送られた注目物体の方位情報を用いて物体の存在領域を推定し、かつ物体方向の平面透視投影画像を提示することにより注目物体の確認が可能であった。また、受信したい全方位動画像を監視者が手動で選択しているため、3視点から自由に監視を行うことができる。したがって、ビルなどの多くのセンサを必要とする場合、3台のサーバ(全方位画像センサ)が同一直線上に並んでいなく、各センサから他の2センサが写っているという前提を満たしながら、3台の全方位画像センサの組み合わせを変化させることで、複数の部屋の監視に適用可能であり、一極集中型の監視を行うことができると考える。

さらに有効なテレプレゼンスシステムと遠隔監視システムを構築するために、今後の課題としては、以下の6つが挙げられる。

## [1] 両システムの共通課題

1. 提示画像の解像度の向上.
2. 使用する帯域幅の削減.
3. 複数の受信(クライアント)側に全方位動画像を伝送するマルチキャスト型システムへの拡張.

1.の問題に関しては,特に全方位画像センサの双曲面の中心付近に投影される領域で解像度が低い.これは,HyperOmni Visionの本質的な特性であり,異なった光学特性に基づいたより高精細全方位画像センサを使用することで解像度の向上が見込まれる.また,2.の問題に関しては,特にテレプレゼンスにおいて,30fpsで全方位画像を伝送するための必要な帯域幅が大きく,ADSLや10Mbps Ethernet,11Mbps無線LANのような一般的なネットワークでは,使用する帯域幅を削減しなければならない.しかし,使用する帯域幅を削減し,全方位画像の伝送フレームレートを抑えると,使用者に違和感を与えてしまう.このような問題に対しては,Windows MediaやMPEGのような,さらに圧縮率の高い伝送方法を用いた帯域幅の削減が有効であると考えられる.また,3.の問題に関しては,動画像をマルチキャストすることによりシステムの拡張が可能であると考えられる.

## [2] 遠隔監視システムの課題

1. フレーム間での存在領域の対応付けによる同一人物の追跡.
2. 壁や窓などへの物体の写りこみを考慮した注目物体の検出.
3. 提示画像における滑らかな仮想カメラワークの実現

1.の問題に関しては,フレーム間で物体の存在領域が未対応であり,受信する全方位画像における検出された物体の画素数を基準に透視投影画像の提示位置を決定している.その結果,物体の画素数により表示位置が変更される場合が多く,監視者にとってわかりづらい.そのため,フレーム間で存在領域を対応付け,同じ表示位置に同一物体の画像を提示し続ける必要がある.また,2.の問題に関しては,一般のカメラにおける物体検出において写り込みや影を分離する手法が提案されて

おり、それらの手法を応用することなどが考えられる。3. に関しては、クライアント側において、カルマンフィルタ等を用いることにより現フレームにおける注目物体の方位情報を推定し、提示画像において滑らかな仮想カメラワークが可能であると考える。

なお、これらの課題を考慮して両システムの実利用を考えた場合には、ネットワークを利用した画像配信に対する要求事項として、以下のことが挙げられる。

1. 動画像のマルチキャスト。
2. 狭帯域での動画像配信。
3. 高品質動画像の伝送。
4. 時間遅延のない動画像配信。

上記 1. の要求事項を満たした動画像配信を行うことが可能であれば、複数のユーザ側への同時動画像配信を行うことが可能となる。またそれに加え 2. の要求事項が可能であれば、複数の全方位動画像の受信を行うことが可能となり、任意地点間での同時テレプレゼンスや一極集中型同時遠隔監視を行うことが可能となる。また無線ネットワークでの動画像配信を行うことで、移動ロボットや車載型テレプレゼンスを行うことも可能となる。3. の要求事項を満たした動画像配信を行うことで、より高解像度の動画像を伝送することが可能となり、さらに高い没入感をユーザに与えることが可能となる。特に遠隔監視では、高解像度の画像を取得することで、移動物体の顔や服の色などの特徴量を用いることができ、より高精度の物体検出や同一物体の追跡を行うことが可能となる。また、4. で示したような監視環境における画像取得から監視者側における画像表示までの処理を時間遅延なく行うことは、セキュリティの上で非常に重要なことであると考えられる。

さらに、テレプレゼンスや遠隔監視などの使用目的や利用可能なネットワークの帯域幅に応じて、上記に述べたこれら動画像の品質保持を重視した動画像配信や時間遅延のない動画像配信を選択することで、ユーザのポリシーに応じたシステム構築が可能であると考えられる。

## 謝辞

本研究を進めるにあたり、その全過程において暖かく御指導、御鞭撻を頂いた視覚情報メディア講座 横矢 直和教授に心より感謝申し上げます。

また、副指導教官として御指導、御鞭撻を頂いた情報科学センター 砂原 秀樹教授、ならびに本研究を進めるにあたり、御助言や研究環境の整備等で多大な御協力を頂いた視覚情報メディア講座 山澤 一誠助教授に心より感謝申し上げます。ならびに公私共々、大変御世話になった視覚情報メディア講座 神原 誠之助手に心より感謝申し上げます。

さらに、公私にわたり多くの御助言を頂いた大阪大学 サイバーメディアセンター 町田 貴史助手、本研究に対する適切な御助言を頂いた視覚情報メディア講座 佐藤智和氏、二年間苦楽を共に過ごしてきた 池田 聖氏、江藤 誠彦氏、小田島 太郎氏、奥村 文洋氏、天目 隆平氏、福富 弘敦氏、英語を御教授して頂いた Steve Vallerand 氏、学部時代より公私に渡って支えあってきた音情報処理学講座 岩見 洋平氏、白石 達也氏、山出 慎吾氏、神戸大学 工学部 川津 義人氏、羽瀧 祐介氏、深井 一成氏、山内 崇氏、山形 太一氏、和木 将晃氏に深く感謝致します。そして一年間ではありますが本研究における実験や機材の手伝いなど様々な面で支えて頂いた 浅井 俊弘氏、峠田 正樹氏、牧田 孝嗣氏に心より感謝の意を表します。また、日々の研究室での生活を支えて頂いた視覚情報メディア講座、事務補佐員 北川 知代女史に心より感謝致します。さらに、今日までの生活において心身ともに気遣って下さった父 森田 一二、母 森田 陽代に心より深く感謝の意を表します。最後に、二年間の研究生活において、たくさんの励ましを下さった中川 美樹女史に心から感謝の意を表します。

## 参考文献

- [1] "Special issue on immersive telepresence," IEEE Multi Media, vol.4, no.1, pp.17-56, 1997.
- [2] S.E. Chen: "QuickTime VR - An Image Based Approach to Virtual Environment Navigation", Proc. of SIGGRAPH 95, pp.29-38, 1995.
- [3] M. Hirose, K. Yokoyama, S. Sato: "Transmission of Realistic Sensation: Development of a Virtual Dome", Proc. of Virtual Reality Annual International Symposium, pp.125-131, 1993
- [4] 山澤 一誠, 八木 康史, 谷内田 正彦: "移動ロボットのための全方位視覚センサ HyperOmni Vision の提案," 信学論, Vol.J79-D-II, No.5, pp.698-707, 1996.
- [5] 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.399-406, 1998.
- [6] T. Uhlin, P. Nordlund, A. Maki and J.O. Eklundh: "Towards an active visual observer," Proc. of ICCV'95, pp.679-686, 1995.
- [7] S. Rougeaux and Y. Kuniyoshi: "Velocity and disparity cue for robust real-time binocular tracking," Proc. of CVPR'97, pp.1-6, 1997.
- [8] 村瀬 健太郎, 和田 俊和, 松山 隆司: "回転を伴うカメラによる移動物体の検出," 画像の認識・理解シンポジウム (MIRU'98), Vol.I, pp.425-430, 1998.
- [9] Q. Cai and J.K. Aggarwal: "Tracking human motion using multiple cameras," Proc. of ICPR'96, pp.68-72, 1996.
- [10] T.J.Olson and F.Z.Brill: "Moving object detection and event recognition algorithms for smart camras," DARPA Image Understanding Workshop, Vol.1, pp.159-175, 1997.

- [11] T.Mori, Y.Kamisuwa, H.Mizoguchi, and T.Sato: “Action recognition system based on human finder and human tracker,” Conf. of IROS’97, pp.1334–1341, 1997.
- [12] T.Kanade, T.Collins, A.J.Lipton, P.Anandan, P.Burt, and L.Wixson: “Co-operative multi-sensor video surveillance,” DARPA Image Understanding Workshop, Vol.1, pp.115–122, 1998.
- [13] 中澤 篤志, 加藤 博一, 井口 征士: “分散カメラシステムによる人物の追跡,” 画像の認識・理解シンポジウム (MIRU’98), Vol.II, pp.1–6, 1998.
- [14] 十河 卓司, 石黒 浩, M.M. Trivedi: “複数の全方位画像センサによる実時間人間追跡システム,” 信学論, Vol.J83-D-II, No.12, pp.2567–2577, 2000.
- [15] 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,” 信学技報, PRMU2000–195, 2001.
- [16] Q. Huang, Y. Cui, S. Samarasekera, and M. Greiffenhagen: “Auto cameraman via collaborative sensing agents,” Proc. of ACCV’98, pp.418–425, 1998.
- [17] 金春 利幸, 八木 康史, 矢内田 正彦: “全方位視ロボットと両眼視ロボットの連携による人物の発見と注視,” 画像の認識・理解シンポジウム (MIRU’98), Vol.II, pp.7–12, 1998.
- [18] D. Gutchess and A. K. Jain: “Automatic surveillance using omnidirectional and active cameras,” Proc. of ACCV2000, pp.228–233, 2000.
- [19] “Combined Omnidirectional and Pan/Tilt/Zoom System,” Research Project in Omnidirectional Surveillance,  
<http://www.cs.columbia.edu/CAVE/VSAM/projects/>.

- [20] 森田 真司, 山澤 一誠, 横矢 直和: “全方位画像センサによるネットワークを介した遠隔監視システム,” 情報科学技術フォーラム (FIT), Vol.3, No.I-38, 2002.
- [21] 森田 真司, 山澤 一誠, 横矢 直和: “全方位画像センサによるネットワークを介したテレプレゼンスシステムと遠隔監視システム,” 信学技報, PRMU2002-83, 2002.
- [22] 森田 真司, 山澤 一誠, 横矢 直和: “全方位カメラによるネットワークを介した遠隔監視システム,” 電気関係学会関西支部連合大会講演論文集, No.G14-11, 2002.
- [23] S. Morita, K. Yamazawa, and N. Yokoya: “Internet Telepresence by Real-Time View-Dependent Image Generation with Omnidirectional Video Camera,” Proc. 1st Int. Workshop on Man-machine Symbiotic Systems, No.DI-8, pp.8, 2002.
- [24] S. Morita, K. Yamazawa, and N. Yokoya: “Internet Telepresence by Real-Time View-Dependent Image Generation with Omnidirectional Video Camera,” Proc. SPIE Electronic Imaging, Vol.5018, pp.51-60, 2003.
- [25] A.Ogawa, K.Kobayashi, O.Nakamura, and J.Murai: “Design and Implementation of DV Stream Over Internet”, IWS Internet Workshop, 1999.
- [26] G.Wolberg: Digital Image Warping, IEEE Computer Society Press, 1990.
- [27] 長井 敦, 久野 義徳, 白井 良明: “時空間情報に基づく侵入者監視システム,” 信学技報, PRU95-204, 1996.
- [28] 二宮 芳樹, 太田 充彦: “オプティカルフローによる移動物体の検知,” 信学技報, PRMU97-28, 1997.
- [29] D.M. Gavrila and V. Philomin: “Real-time object detection for ”smart” vehicles,” Proc. of ICCV'99, pp.87-93, 1996.

- [30] 村上 卓也, 菊地 仁一: “画像処理を用いた侵入者検知装置 – MP30,” 信学技報, IE95-51, 1995.
- [31] 三木 隆太郎, 山澤 一誠, 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,” 信学技報, PRMU98-177, 1999.

## 研究業績

1. 森田 真司, 山澤 一誠, 横矢 直和: “全方位画像センサによるネットワークを介した遠隔監視システム,” 情報科学技術フォーラム (FIT), Vol.3, No.I-38, 2002.
2. 森田 真司, 山澤 一誠, 横矢 直和: “全方位画像センサによるネットワークを介したテレプレゼンスシステムと遠隔監視システム,” 信学技報, PRMU2002-83, 2002.
3. 森田 真司, 山澤 一誠, 横矢 直和: “全方位カメラによるネットワークを介した遠隔監視システム,” 電気関係学会関西支部連合大会講演論文集, No.G14-11, 2002.
4. S. Morita, K. Yamazawa, and N. Yokoya: “Internet Telepresence by Real-Time View-Dependent Image Generation with Omnidirectional Video Camera,” Proc. 1st Int. Workshop on Man-machine Symbiotic Systems, No.DI-8, pp.8, 2002.
5. S. Morita, K. Yamazawa, and N. Yokoya: “Internet Telepresence by Real-Time View-Dependent Image Generation with Omnidirectional Video Camera,” Proc. SPIE Electronic Imaging, Vol.5018, pp.51-60, 2003.