

# 第5章 視点変換アルゴリズムを統合した奥行き獲得手法

本章では HMD を装着した観察者視点からの奥行きマップのように、奥行きマップを獲得したい位置に高精度な奥行きセンサを設置することが困難な場合の奥行き獲得法について提案する。本研究では、奥行きマップが必要な視点とは異なる視点に設置した奥行きセンサから獲得した奥行きマップを視点変換手法を用いて奥行きマップが必要となる視点に変換する、クライアント・サーバ型の奥行きセンシング方式を提案し、奥行きセンサを装着しない観察者の自由視点での奥行きマップの獲得を実現する。また、提案手法をサーバセンサとしてビデオレートのステレオセンサ、クライアントセンサとして赤外線センサにより構成し、視点変換によって獲得した観察者視点の奥行きマップをもとに現実物体と仮想物体の間の隠れを正しく再現した CG 合成を行うシステムを構築する。

## 5.1 観察者視点の奥行き獲得のためのセンシング方式

### 5.1.1 観察者視点の奥行き獲得の問題点

複合現実感において、現実世界と仮想世界の違和感のない融合を実現するためには、「幾何的整合性」、「光学的整合性」、「時間的整合性」を解決する必要がある。幾何的整合性には、現実世界と仮想物体の位置や姿勢を正しく整合させる問題、及び、実物体と仮想物体の前後関係による隠れや現実世界に投影される仮想物体の影を表現する問題が含まれる。後者を正しく表現しないと図 2.9(a) に示したように、位置や姿勢を正しく表現しても仮想物体の 3 次元位置の把握が著しく困難な融合結果になる。実物体によって隠される仮想物体を表現するためには、図 2.9(b) のように隠される部分を消去して描画する必要があり、隠されるか否かを画素ごとに判断するために、観察者、すなわち、観察者が装着している HMD の視点からの現実世界の密な奥行きマップが必要になる。

多視点ステレオ法を用いて観察者視点の奥行き情報をビデオレートで獲得し、実物体と仮想物体の合成を、隠れを考慮して実現した例として金出らの Z-key[16] がある。しかし、多視点ステレオ画像を撮影するための多眼カメラヘッドは大きくて重い。超小型カメラを用いれば軽いカメラヘッドを製作することも可能かもしれないが、カメラヘッドの大きさをおさえるためにカメラ間隔を小さくしすぎると、得られる奥行き情報の精度が低下するというジレンマに陥る。

複合現実感において、シースルー型 HMD を装着した観察者は現実世界を移動することが前提であり、観察者視点の奥行きマップを得るために、HMD に多眼ステレオヘッドを装着する必要がある。しかしこれは、大きさや重さの点で非現実的である。多視点ステレオ法だけでなく、レーザレンジファインダ法を見渡しても、HMD に装着可能な程度に小型軽量で、ビデオレートに近い速度で密な奥行きマップを獲得可能なセンサを実現することは、現時点では不可能に近いと考えられる。

### 5.1.2 クライアント・サーバ方式による観察者視点の奥行き獲得

このようなジレンマを回避し、動き回る観察者が装着した HMD の視点からの現実世界の奥行き情報を獲得できる手法として、クライアント・サーバ方式による奥行きセンシング法を提案する。これは、観察者の近傍にサーバセンサとして奥行きセンサを固定して設置し、それによって得られる現実世界の奥行き情報を、観察者に装着されたセンサから得られる位置・姿勢情報にもとづいて、観察者の視点からの奥行き情報として視点変換し供給するものである。図 5.1 にクライアント・サーバ方式を用いた奥行き獲得手法の概念図を示す。

クライアント・サーバ型の奥行きセンシング方式を採用することで以下に挙げるような利点が得られる。

1. 観察者の HMD に装着するセンサは基本的に位置センサで良いため、小型・軽量化できる。
2. サーバセンサは固定して設置するため、大型のセンサを利用することができます。また、異なる種類の複数のセンサを組み合わせることも可能である。これにより、精度の良い奥行き獲得が実現される。
3. サーバセンサを実世界に固定することにより、サーバセンサから観測されるシーン中のほとんどの部分を静止しているものと考えられる。これにより、奥行き獲得精度の向上が期待できる。更に、静止している対象については、奥行き獲得に多少時間はかかるが精密な奥行きが獲得可能なレーザレンジファインダを利用することができます。
4. 複数のサーバセンサを用いて奥行き情報を獲得することができる。人などの動物体に対しては、ビデオレートステレオマシンを利用する。静止物体については、レーザレンジファインダを用いて正確な奥行きを獲得する。また、壁や床、大きな家具といった固定されて動かないものに対しては、予め 3 次元のポリゴンモデルを用意しておくことも可能である。
5. サーバセンサで獲得した奥行き情報を複数の観察者で共有することができる。これは、観察者一人一人に個別のセンサを設定する方式と比べてコストの面でも有利である。

## 5.1 クライアント・サーバ型の奥行きセンシング

この章では、クライアント・サーバ型の奥行きセンシングについて述べる。

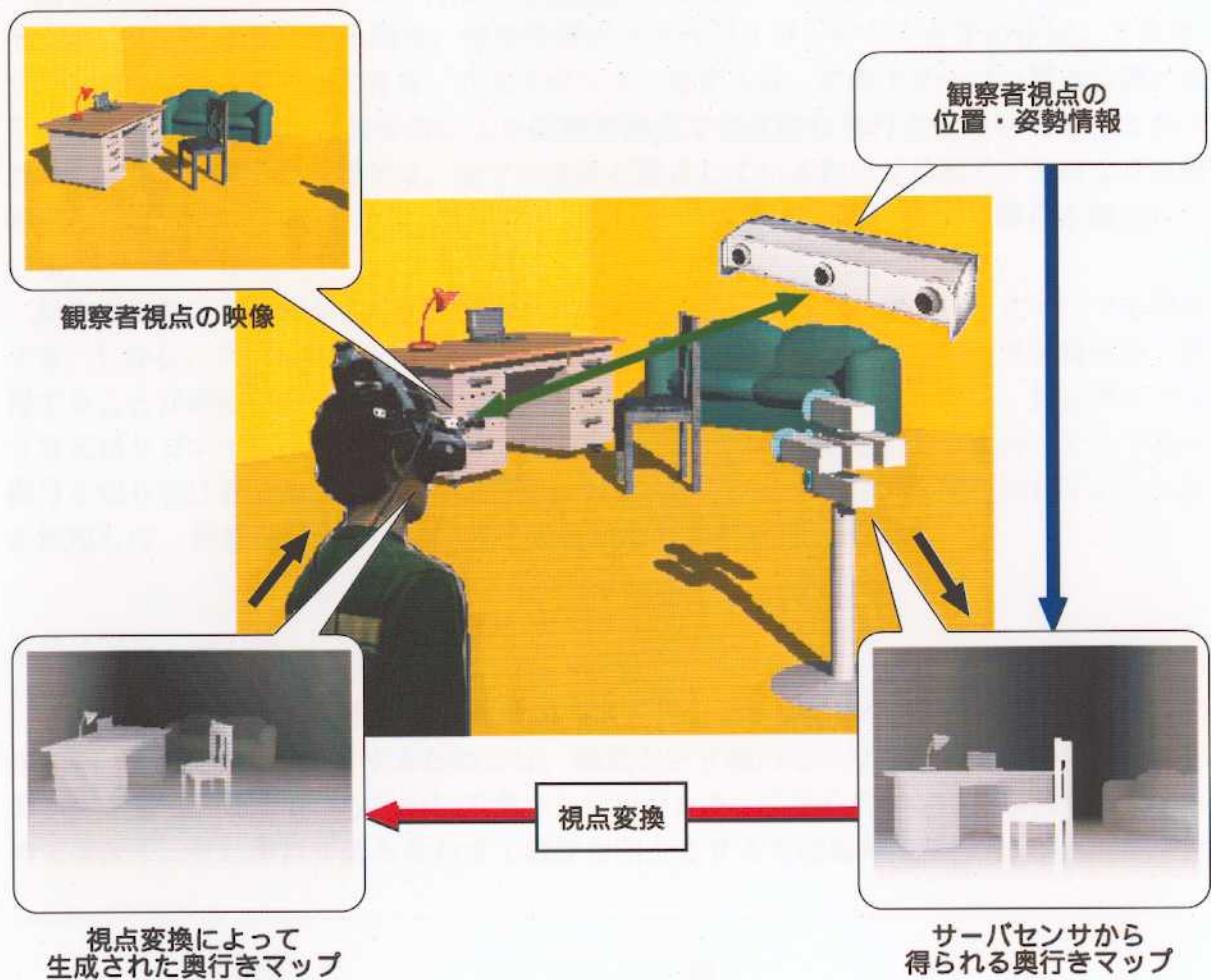


図 5.1: クライアント・サーバ型の奥行きセンシング方式の概要

### 5.1.3 クライアントとサーバセンサの構成

本節では、クライアント・サーバ型のセンシング方式を実現するための各センサの構成について考察する。

#### サーバセンサ

図 5.2 にサーバセンサとして利用可能な奥行き獲得センサの例を挙げる。現実世界を完全な静止世界と仮定できる場合、現実世界の 3 次元ポリゴンモデルを予め作成しておき、サーバセンサとして利用できる。3 次元ポリゴンモデルは、正確でサーバの視点位置に依存しないため、簡単な幾何変換により観察者視点での正確な奥行き情報を得ることができる。しかし、現実の世界では、全ての物体が静止しているわけではなく、人のような動物体も存在する。本研究でも、このようなシーン中に動物体が存在する場合を想定している。

動物体の奥行き情報を実時間で獲得するために、ビデオレートのステレオセンサを利用する。しかし、ステレオセンサではテクスチャの少ない滑らかな面の奥行きを精度良く獲得することが困難である。そこで、動きのある部分ではステレオセンサを、静止部分では 3 次元ポリゴンモデルを組み合わせて利用する。この場合、動きのある部分とそうでない部分を切り分ける必要がある。この処理はステレオセンサが実世界に固定されていることをを利用して、背景差分法により用意に実現することが可能である。

#### クライアントセンサ

クライアントセンサで獲得すべき情報は観察者の装着 HMD の位置・姿勢情報である。HMD の位置・姿勢を獲得するためには、磁気センサ等のセンサを用いる方法とシーン中のランドマークを画像から抽出して求める方法がある。これらのセンサを単独で使用するのではなく、それぞれを組み合わせて精度を向上させる方法も考えられる。図 5.3 にクラ



(a) ビデオレートステレオマシン



(b) レーザレンジファインダ



(c) 3 次元ポリゴンモデル

図 5.2: サーバセンサ



図 5.3: クライアントセンサ

イアントセンサとして利用可能な位置センサを挙げる。

本研究では、サーバセンサとして動領域に対してビデオレートのステレオセンサ、静止領域に対して3次元ポリゴンモデル、また、クライアントセンサとして赤外線センサを用いたクライアント・サーバ型の奥行きセンシング方式の実現を目指す。以下では、ステレオセンサから獲得した奥行きマップを視点変換手法を用いて異なる視点からの奥行きマップに変換する手法、視点変換によって得た奥行きマップの高精度化についての検討する。

## 5.2 視点変換アルゴリズムによる奥行きマップ生成

本節では、ステレオセンサで獲得された奥行きマップをクライアントセンサから得られるHMDの位置・姿勢情報をもとに視点変換することにより、観察者視点での奥行きマップを生成する手法について説明する。

### 5.2.1 視点変換による別視点での奥行きマップ生成

図5.4に各センサおよび現実世界の座標系の関係を示す。 $(x_c, y_c, z_c)$ はクライアントセンサの位置を原点とした座標系、 $(u, v)$ はクライアントカメラで撮影される2次元画像の座標である。また、 $(x_s, y_s, z_s)$ はサーバセンサで獲得される奥行き情報を表現する座標、 $(x, y, z)$ は現実世界での物体の位置、すなわちワールド座標系を表す。

この時、ワールド座標系でのサーバセンサの位置・姿勢が与えられると、サーバセンサで獲得した奥行き情報  $p_s = [x_s, y_s, z_s]^T$  と対象物体のワールド座標系での位置  $p_w = [x, y, z]^T$  との関係を式(5.1)で表現することができる。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (5.1)$$

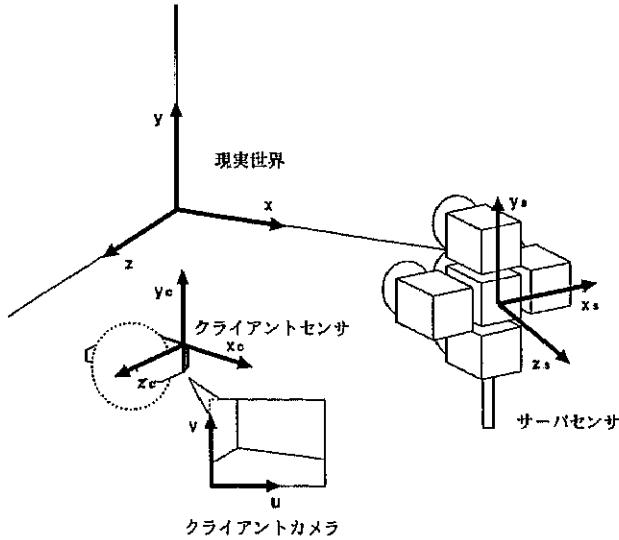


図 5.4: 各センサで用いられる座標系

このとき、 $T_s = [T_x, T_y, T_z]^T$  はワールド座標系でのサーバセンサの位置、 $r_{ij}$  はサーバセンサの姿勢を表す回転行列  $R_s$  の成分で、以下の式で与えられる。 $\theta_x, \theta_y, \theta_z$  はそれぞれの軸周りの回転成分を表す。

$$R_s = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.2)$$

また、式 (5.1) は同次座標を用いて式 (5.3) のように表現できる。

$$\tilde{p}_w = M_s \tilde{p}_s \quad (5.3)$$

ここで、 $M_s$  は、

$$M_s = \left[ \begin{array}{c|c} R_s & T_s \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \quad (5.4)$$

で与えられる。また、 $\tilde{p}_w, \tilde{p}_s$  は  $p_w, p_s$  の同次座標表現である。

更に、ワールド座標系でのクライアントセンサの位置・姿勢が与えられると、式 (5.3) と同様にワールド座標系  $p_w$  とクライアント座標系  $p_c$  の関係が得られる。 $M_c$  は式 (5.4) と同様にクライアントセンサの位置・姿勢から決定される変換行列である。

$$\tilde{p}_w = M_c \tilde{p}_c \quad (5.5)$$

この式は  $M_c$  の逆行列  $M_c^{-1}$  を用いて式(5.6)のように表せる。

$$\tilde{p}_c = M_c^{-1} \tilde{p}_w \quad (5.6)$$

以上のことから、式(5.3)(5.5)を用いて、サーバセンサから得られた奥行きマップをクライアントセンサからの奥行きマップへ変換することができる(式(5.7))。

$$\tilde{p}_c = M_c^{-1} M_s \tilde{p}_s \quad (5.7)$$

### 5.2.2 物体輪郭での奥行き情報の高精度化

ステレオセンサで獲得した奥行きマップは、ステレオの対応点探索精度の問題から、本来滑らかに変化するはずの奥行き値が不均一な値になることがある。このような場合、サーバ側で獲得された奥行き値の微妙な間違いが視点変換後の奥行きマップに大きな影響を与えることがある。特に、物体輪郭周辺で奥行き値にはらつきがある場合、視点変換によって得た奥行きマップの物体輪郭は、実際の物体輪郭と大きく異なる。現実物体が仮想物体を隠す表現を忠実に再現するためには、現実物体の輪郭が正確な奥行きマップを獲得することが重要である。ステレオ法の対応点探索に用いられる対応窓の大きさをある程度大きく設定することにより、滑らかな奥行きマップを獲得することが可能であるが、獲得される奥行きマップの輪郭が実際の物体輪郭よりも太ってしまうという問題が生じる。

そこで、ステレオセンサで獲得された奥行きマップを修正する方法を考える。奥行きマップの修正には、大きく分けて二通りの方法が考えられる。一つは、サーバセンサで獲得された奥行きマップを視点変換する前に修正する方法、もう一つは、視点変換によって得られた奥行きマップを修正する方法である。

#### 視点変換前の奥行きマップの修正

視点変換後の奥行きマップの物体輪郭が不正確になる原因是、獲得された奥行きマップの奥行き値の不均一さによるものである。そこで、サーバセンサで獲得された奥行きマップの物体輪郭を均一な奥行き値に修正することで、視点変換後の奥行きマップの物体輪郭の不正確さを軽減できると考えられる。

物体輪郭での奥行きマップの修正法の例を以下に挙げる。

- 物体全体の奥行き値を一定の値にする

対象が比較的小さい場合もしくは、物体全体で奥行きの変化が小さい場合には有効であるが、同一物体内で奥行き値が大きく変化する場合には、物体輪郭の不一致が生じることがある。

- 物体輪郭の奥行きに対して、平均値フィルタやメディアンフィルタ等の平滑化フィルタを適用する

物体輪郭の奥行き値が比較的滑らかである場合、平滑化によって更に周囲の奥行き値との変化が滑らかになり、視点変換によって得られる奥行きマップの物体輪郭も滑らかになると考えられる。

### 視点変換後の奥行きマップの修正

前述の視点変換前の奥行きマップの修正が有効に働くためには、サーバセンサで獲得される奥行きマップの物体輪郭が実際の物体輪郭と一致しているという前提条件が必要である。ステレオ法によって獲得される奥行きマップでは、物体輪郭が実際の輪郭よりも外側に広がる傾向にあるため、視点変換前の奥行きマップを修正することで、輪郭位置の正確な滑らかな奥行きマップを得ることは困難である。

最終的に観察者視点からの実際の物体輪郭と一致した滑らかな奥行きマップが得ることが目的であるから、視点変換後の奥行きマップの輪郭を修正する方が有効であると考えられる。本研究では、観察者のHMDに装着されたCCDカメラによって得られた観察者視点映像から物体の輪郭を抽出し、抽出した輪郭情報をもとに、視点変換によって得られた奥行きマップの輪郭を修正する方法を試みる。輪郭修正の手順は以下の通りである。

1. 観察者視点映像に対してエッジ検出オペレータを適用して物体輪郭を抽出する
2. 視点変換によって得られた奥行きマップの輪郭を以下の手順に従って抽出する(図5.5)
  - (a) 奥行き画像を左から物体が見つかるまで走査する(物体以外の領域は奥行き値0を持つ)
  - (b) 見つかった点を物体の左輪郭とする
  - (c) 更に画像を走査し、物体でなくなる点を見つける
  - (d) 見つかった点を物体の右輪郭とする
  - (e) 全画面について上記の処理を行う
3. 観察者視点映像から抽出した輪郭情報をもとに奥行きマップの輪郭を修正する(図5.6)
  - (a) 奥行きマップから抽出した左輪郭の位置から観察者視点映像から抽出した物体輪郭画像を左に走査し、輪郭点を探す
  - (b) 輪郭点が見つかった場合、左輪郭の位置から実際の輪郭点までの奥行き値を0とする
  - (c) 輪郭点が見つからなかった場合、左輪郭の位置から観察者視点映像から抽出した物体輪郭画像を右に走査し、輪郭点を探す

- (d) 輪郭点が見つかった場合、左輪郭の位置から輪郭点までの奥行き値を左輪郭の奥行き値で補完する
- (e) 右輪郭点に関して、左輪郭点と画像の走査方向を反転させて (a) から (d) までの処理を行う

この方法は、実際に観察者視点映像から得られる物体輪郭をもとに奥行きマップの輪郭を修正するため、実際の物体輪郭に忠実な滑らかな奥行きマップを得ることができる。ここで、得られる奥行きマップの輪郭の精度は観察者視点映像から抽出される物体輪郭の精度に依存する。精度の良く奥行きマップの輪郭を修正するためには、物体輪郭の抽出精度の向上が望まれる。

## 5.3 実験結果

### 5.3.1 視点変換による奥行きマップ生成

5.2.1 節で述べた視点変換手法を用いてサーバセンサから得られた奥行きマップを観察者視点からの奥行きマップに変換する実験を行った。サーバセンサの位置・姿勢は、予めワールド座標系での3次元座標が既知なマーカを対象となる空間に配置しておき、マーカの3次元座標とステレオセンサの中心カメラで撮影した画像中のマーカの画像座標との対応関係から算出しておく。また、クライアントセンサには赤外線センサを利用し、観察者が装着しているHMDの位置・姿勢を獲得する。

図5.7に実験結果を示す。図5.7(a)はサーバセンサで獲得された映像、図5.7(b)は観察者が装着しているHMDに取りつけられたCCDカメラから得られた映像である。図5.7(c)(d)はサーバセンサで獲得されたシーンの奥行きマップである。図5.7(d)は動物体を含まない状態で予め獲得しておいた奥行きマップで、この奥行きマップと動物体が入ってきた時の奥行きマップとの差分を取ることで、図5.7(e)に示すような動物体のみの奥行

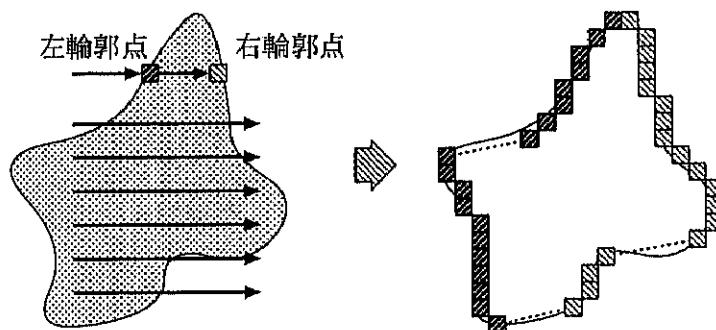


図 5.5: 奥行きマップからの輪郭抽出

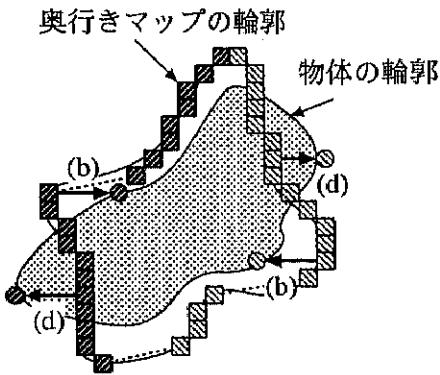


図 5.6: 奥行きマップの輪郭修正

きマップを得ることができる。この奥行きマップを視点変換手法により観察者視点からの奥行きマップに変換したのが図 5.7(f) の奥行きマップである。

### 5.3.2 奥行きマップの輪郭修正

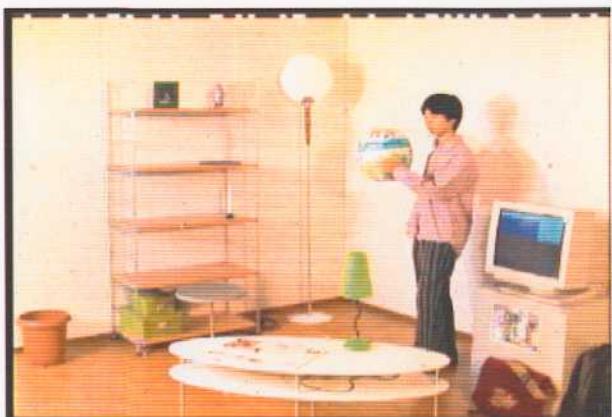
5.2.2 節で述べた手法に従って、奥行きマップの輪郭修正の実験を行った。視点変換による奥行きマップの獲得は前節と同じ条件で行った。実験結果を図 5.8 に示す。図 5.8(a) は観察者視点映像である。この映像に対してラプラシアンフィルタを適用し、適当な閾値で二値化することで、図 5.8(b) に示す物体輪郭映像を生成する。また、視点変換によって獲得された奥行きマップ図 5.8(c) から抽出した物体輪情報が図 5.8(d) である。図 5.8(b), 5.8(d) をもとに図 5.8(c) の輪郭を修正した奥行きマップが図 5.8(e) である。図 5.8(e) で物体輪郭の奥行きマップが滑らかに修正されていることが確認できる。

## 5.4 奥行き情報共有型の複合現実感システムの構築

### 5.4.1 システムの構成

サーバセンサとしてビデオレートのステレオセンサ(図 5.2(a))と静止物体の 3 次元ポリゴンモデル(図 5.2(c)), クライアントセンサとして赤外線センサ(図 5.3(a))を用いて、視点変換によって獲得した観察者視点の奥行きマップをもとに現実物体と仮想物体の間の隠れを正しく再現した CG 合成を行うシステムを構築した。図 5.9 に構築したシステムの概要を示す。

サーバセンサでは、動物体の奥行きマップをビデオレートのステレオセンサから、静止物体の奥行きマップを予め作成しておいた 3 次元ポリゴンモデルから獲得する。動領域の抽出は、予め獲得しておいた動物体の含まれない奥行きマップと獲得した奥行きマップとの背景差分により行う。ステレオセンサは固定して設置されるため、その位置・姿勢



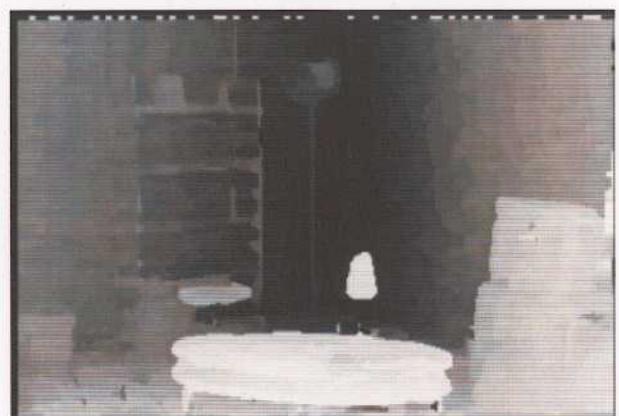
(a) サーバセンサからの映像



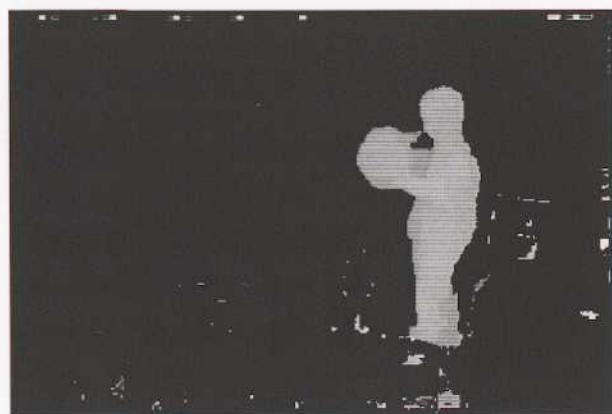
(b) 観察者視点の映像



(c) サーバセンサで獲得した奥行きマップ



(d) サーバセンサで獲得した背景の奥行きマップ



(e) 背景差分により獲得した動領域の奥行きマップ

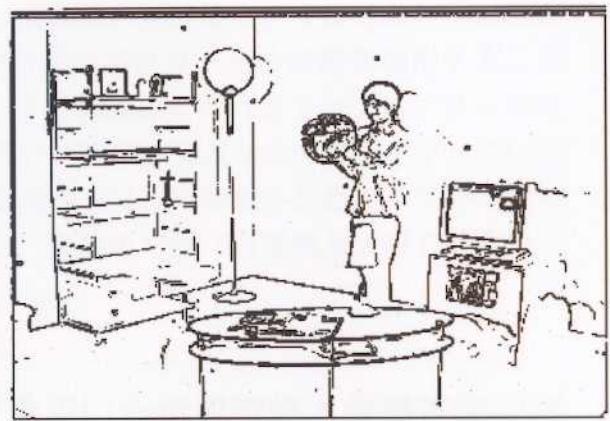


(f) 視点変換によって獲得した動領域の奥行きマップ

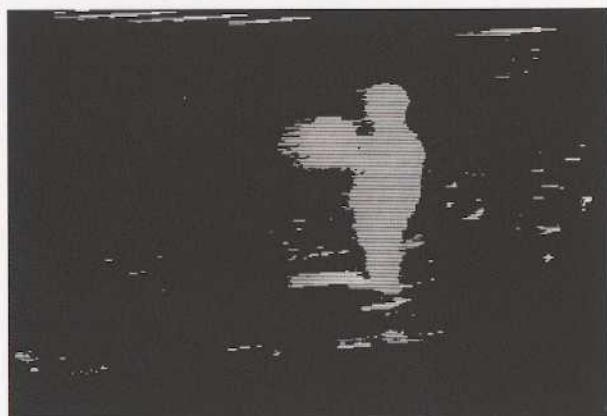
図 5.7: 視点変換手法による奥行き獲得実験結果



(a) 観察者視点映像



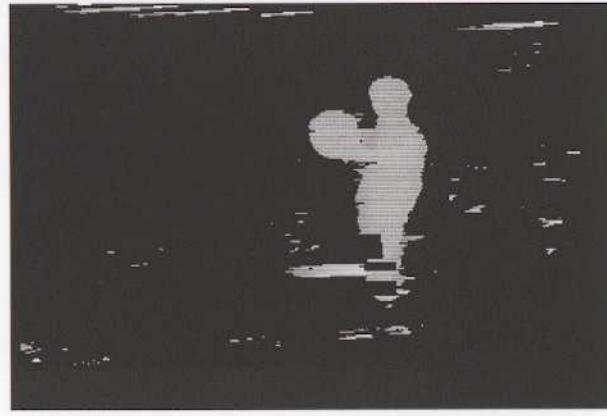
(b) 観察者視点映像から  
抽出した物体輪郭



(c) 輪郭修正前の奥行きマップ



(d) 奥行きマップから  
抽出した物体輪郭



(e) 輪郭修正後の奥行きマップ

図 5.8: 奥行きマップの輪郭修正結果

は、予めキャリブレーションを行うことで獲得しておく。赤外線センサは対象空間に固定して設置され、観察者の装着する HMD に取り付けた LED マーカの位置を検出することで、HMD の位置・姿勢を獲得する。サーバセンサで獲得した奥行きマップは、サーバセンサとクライアントセンサの位置・姿勢情報を用いて観察者視点からの奥行きマップに変換される。変換された奥行きマップをもとに観察者の HMD の装着されたカメラから獲得した観察者視点映像に対して、CG の描画を行う。奥行きマップの変換および CG 合成は SGI Onyx2 ワークステーションによって行う。

#### 5.4.2 観察者視点からの奥行きマップを用いた現実物体と仮想物体の隠れの再現

構築したシステムを用いて観察者視点映像に対して現実物体による仮想物体の隠れを再現した CG 合成を行う実験を行った。ビデオレートステレオマシンからは対象シーンの奥行きが 60 フレーム/秒で獲得される。ステレオセンサで獲得される奥行きマップで視点変換の対象となる領域を動領域に限定することで、高速に観察者視点からの奥行きマップを獲得することが可能となった。また、奥行きマップ生成の際の変換行列による座標計算やテクスチャマッピングにグラフィックボードのジオメトリエンジンを利用することで、視点変換による奥行きマップ生成から CG 合成までの処理速度を 12~15 フレーム/秒で実現することができた。

構築したシステムによる実験結果を図 5.10 に示す。図 5.10(a), (b) はサーバセンサから得られた映像と奥行きマップである。また、図 5.10(c) は予め獲得したおいた動物体を含まない奥行きマップとの背景差分によって獲得した動物体のみの奥行きマップである。観察者の装着した HMD に取りつけられた CCD カメラからは図 5.10(d) に示すような映像が得られる。視点変換手法により得られた観察者視点からの奥行きマップが図 5.10(e) である。これに対して、静止物体の 3 次元ポリゴンモデルから得られる奥行きマップを合成した結果が図 5.10(f) である。最後に、得られた奥行きマップをもとに観察者視点映像に CG を合成することで、図 5.10(g) の結果が得られる。

### 5.5 考察

本章では、HMD を装着した観察者視点からの高精度な奥行きマップを獲得を実現するために、奥行きマップが必要な視点とは異なる視点に設置した奥行きセンサから獲得した奥行きマップを視点変換手法を用いて奥行きマップが必要となる視点に変換する、クライアント・サーバ型の奥行きセンシング方式を提案した。サーバセンサとしてビデオレートのステレオセンサ、クライアントセンサとして赤外線センサを採用し、サーバセンサで獲得した奥行きマップから観察者視点からの奥行きマップを獲得するための視点変換手法について検討した。サーバセンサとして、動物体に対してビデオレートのステレオセンサ、静止物体に対して 3 次元ポリゴンモデルを用いることで、静止物体の奥行きマップは常に

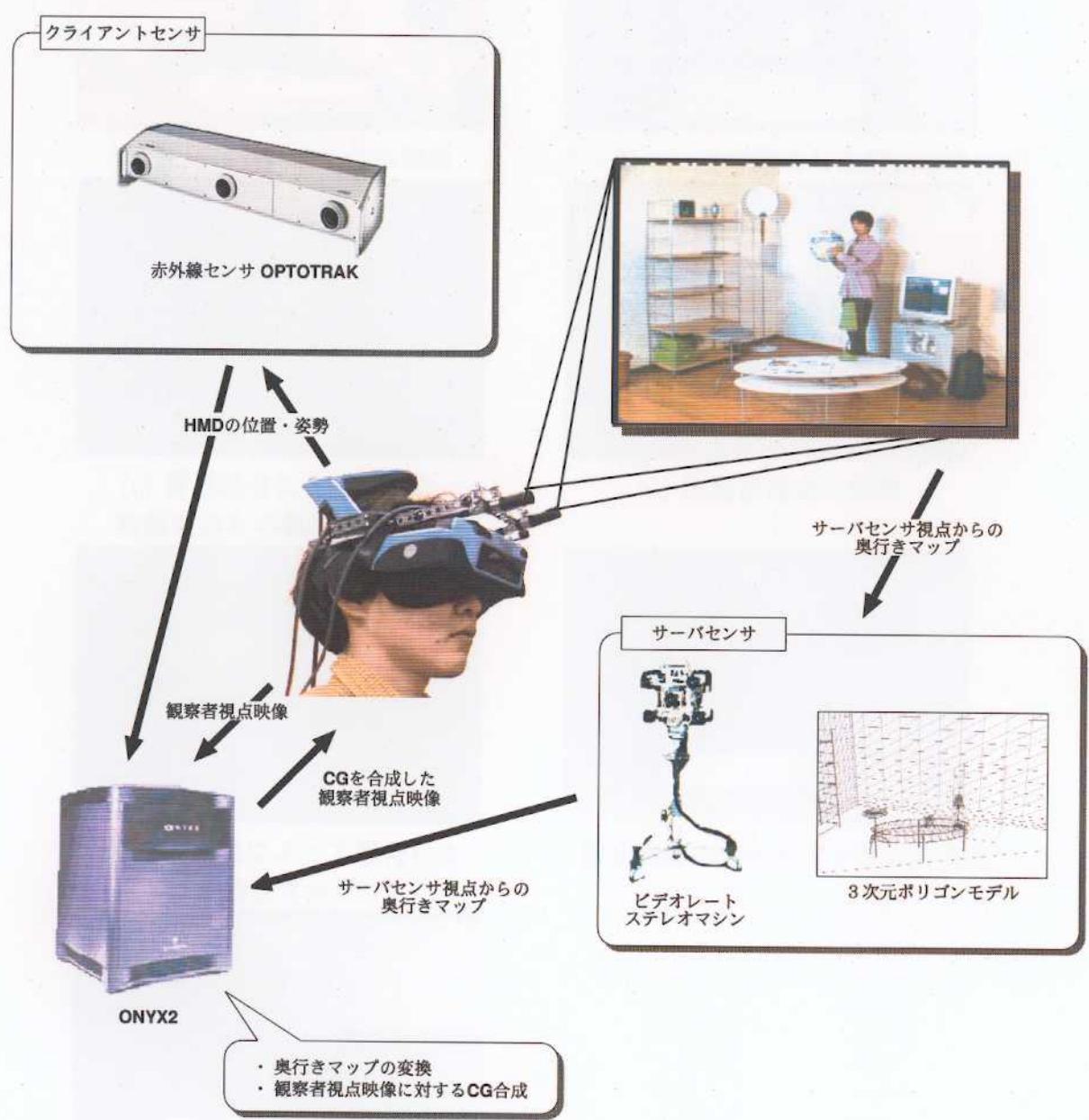


図 5.9: 構築したシステムの概要



(a) サーバセンサからの映像



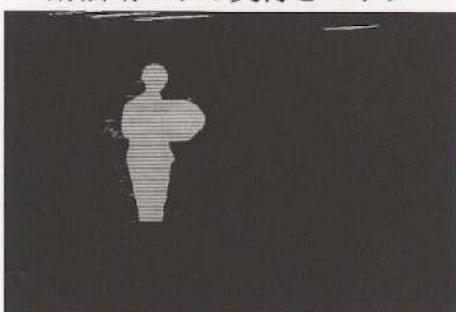
(b) サーバセンサで獲得した奥行きマップ



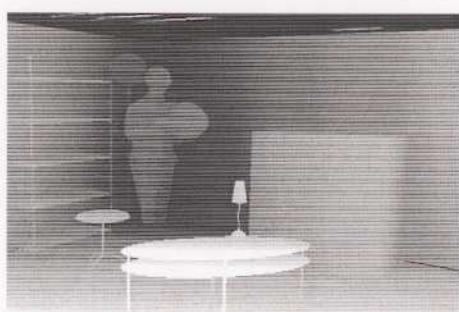
(c) 背景差分によって得た動領域のみの奥行きマップ



(d) 観察者視点の映像



(e) 視点変換によって獲得した奥行きマップ



(f) 静止物体の奥行きマップとの合成



(g) 仮想物体の隠れを再現したCGの合成

図 5.10: 構築したシステムによる実験結果

高精度に、動物体ではビデオレートでの奥行きマップ獲得が実現できることを示した。また、提案した方式を用いた奥行き獲得システムを構築し、視点変換手法によって獲得した観察者視点からの奥行きマップをもとに観察者視点映像に対して現実物体と仮想物体の隠れを再現したCG合成・提示を実現した。今後は、視点変換手法を用いて生成した奥行きマップの輪郭精度を向上させ、観察者に対して自然な映像合成を目指すとともに、システムの高速化について検討する必要がある。