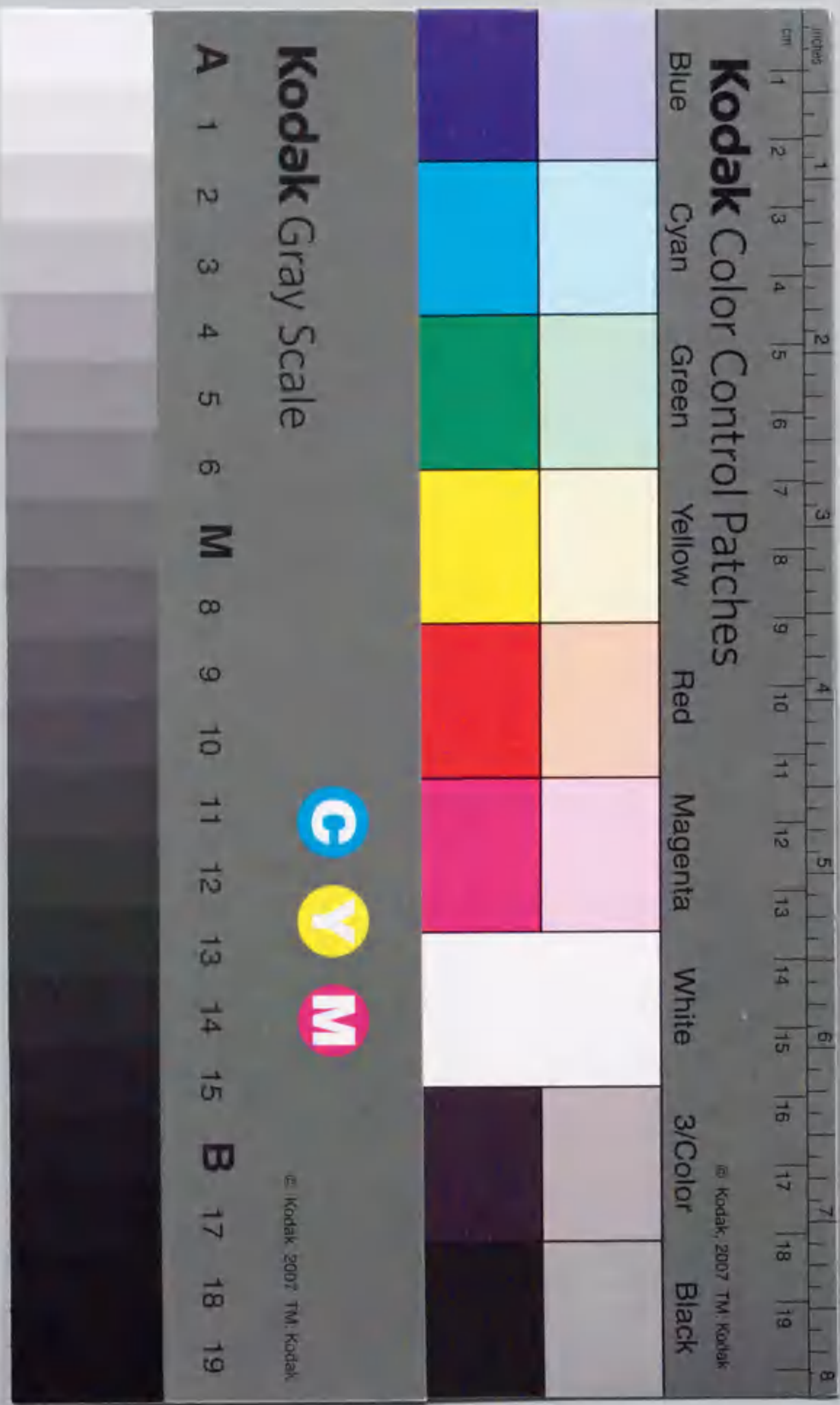


実空間と仮想空間との  
リアルタイム融合に関する研究

Study on Real-time Merging  
of Real and Virtual Spaces

1999年7月

箕浦弘人





①

実空間と仮想空間との  
リアルタイム融合に関する研究

Study on Real-time Merging  
of Real and Virtual Spaces

1999年7月

箕浦 弘人



# 目次

第1章 序論	1
1.1 Virtual Reality	1
1.2 Mixed Reality	2
1.3 Mixed Realityの整合問題	3
1.4 隠蔽関係の表現方法の分類	4
1.5 本論文の目的	6
1.6 本論文の構成	7
第2章 多層分割による隠蔽関係の表現	9
2.1 レイヤー合成の原理	9
2.2 レイヤー合成の表現力	11
2.2.1 垂直分割	11
2.2.2 層の向きによる違い	13
2.2.3 静的な層・動的な層の違い	15
2.2.4 単眼視・立体視の違い	15
2.3 改良方法の提案	17
2.3.1 従来のレイヤー合成システムと改良点	17
2.3.2 提案方法の位置づけ	18
第3章 3層分割を用いた実空間と仮想空間との合成方法	19
3.1 合成方法	19



- 3.2 クリッププレーンの決定 . . . . . 20
- 3.3 システム構成 . . . . . 23
- 3.4 実空間と仮想空間の座標系の整合 . . . . . 28
  - 3.4.1 変換行列の算出 . . . . . 28
  - 3.4.2 GLパラメータとの関係 . . . . . 29
  - 3.4.3 GLパラメータの算出 . . . . . 34
- 3.5 実空間と仮想空間の変換のキャリブレーション . . . . . 40
  - 3.5.1 キャリブレーション方法 . . . . . 40
  - 3.5.2 手作業によるキャリブレーション . . . . . 41
  - 3.5.3 キャリブレーションの自動化 . . . . . 42
  - 3.5.4 自動ポインティング . . . . . 42
  - 3.5.5 多点でのキャリブレーション . . . . . 47
  - 3.5.6 精度の検証 . . . . . 48
  - 3.5.7 左右両眼映像の自動切り替え . . . . . 48
- 第4章 実空間と仮想空間の合成例 . . . . . 51
  - 4.1 必要な層の削減例: バーチャルスタジオ . . . . . 51
    - 4.1.1 システム構成 . . . . . 51
    - 4.1.2 CCDカメラによる位置検出 . . . . . 52
    - 4.1.3 合成結果 . . . . . 57
  - 4.2 連続した物体の合成例: 手を用いた組立作業 . . . . . 60
    - 4.2.1 磁気位置センサによる位置検出 . . . . . 60
    - 4.2.2 ビデオシースルー型HMDを用いた結果提示 . . . . . 61
    - 4.2.3 合成結果 . . . . . 61
- 第5章 結論 . . . . . 63

# 第1章 序論

## 1.1 Virtual Reality

近年 Virtual Reality に関する研究が盛んに行われている。なぜなら Virtual Reality の技術は、21 世紀に向けて、産業・医療・アミューズメントなど多くの分野での実用化が期待されているからである [1][2][3].

この Virtual Reality の技術・システムにおいて特に重要な要素は、D. Zeltzer によると「Autonomy」(自立性)・「Interaction」(対話性)・「Presence」(臨場感)の3要素(図 1.1)にまとめられる [4][5].

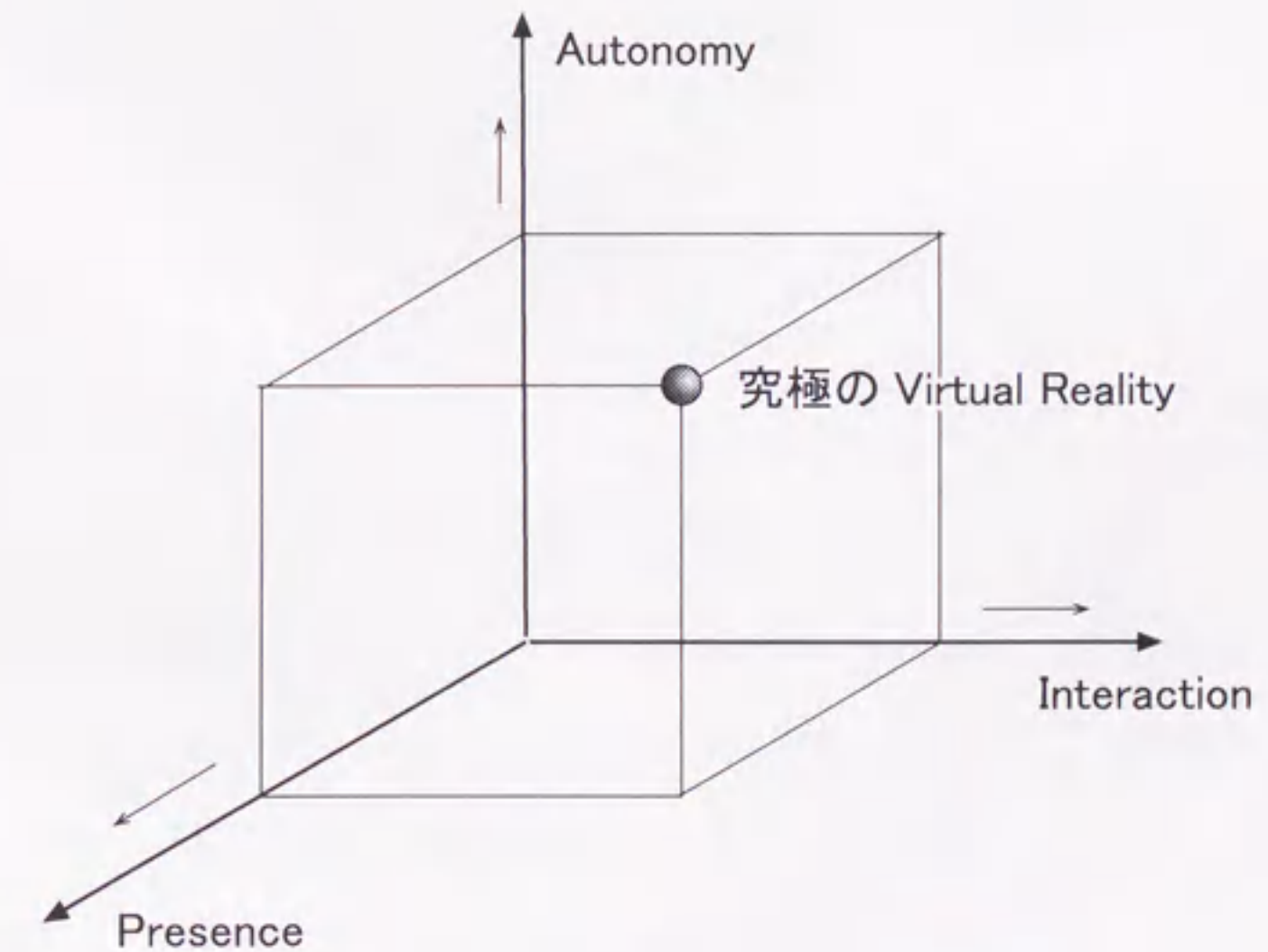


図 1.1: Autonomy, Interaction and Presence.



「Autonomy」は、実空間と同じように仮想空間でも物が重力に引かれて落ちたり、人を見て犬が吠えたりなど、人とは独立して運動・動作することである。物理現象のシミュレータなどは「Autonomy」の要素が強い Virtual Reality である。

「Interaction」は、自分の居る実空間と仮想空間同士で片方の空間の振る舞いが他方の空間に影響を及ぼすことで、両空間が独立して存在するのではなくお互いに空間の情報を共有しながらやりとりすることである。「対話性」を実現している装置の例としては、ペン型の入力・力覚出力をもった装置である PHANTOM[6]などが開発されている。

「Presence」は、恰も本物であるかのような仮想空間(コンピュータグラフィック、ジオラマ映像・合成音など)を作成し提示するだけでなく、自分の空間と作り出された仮想空間の境界を取り外し、深い没入感を得ることである。視覚的装置としては、Head Mounted Display[7][8]やCAVE[9]などといったハードウェアが開発され、CGのリアリティと没入感が得られるようになりつつある。

Virtual Reality 技術・システムの研究はこれら3要素を高めることで、基礎研究やアプリケーションの開発にソフトウェア・ハードウェアの両面からアプローチされている[10][11]。

## 1.2 Mixed Reality

Virtual Reality の一分野であり、実空間と仮想空間の融合をテーマとする Mixed Reality は、Virtual Reality の3要素の中では「Presence」の要素が強いが、実空間と仮想空間の融合という点から3要素すべてに深く関連する重要な分野である。

Mixed Reality は、他に Augmented Reality や Augmented Virtuality とも呼ばれることがあるが、P. Milgram は図 1.2 に示すように、これらを Mixed Reality の「Virtuality Continuum」としてまとめている[12][13]。「Virtuality Continuum」では、実空間と仮想空間を合成する時、どちらをベースとしてどちらを融合するかの割合によって呼ばれ方が違うが、それらに境界はなく連続しており、その continuum を Mixed Reality と呼んでいる。

この実空間と仮想空間の融合とは、視覚・聴覚・聴覚など人間の実際の感覚に、人工感覚をあたかも実際の感覚であるかのように織り交ぜることである。特に人間の情報取得に

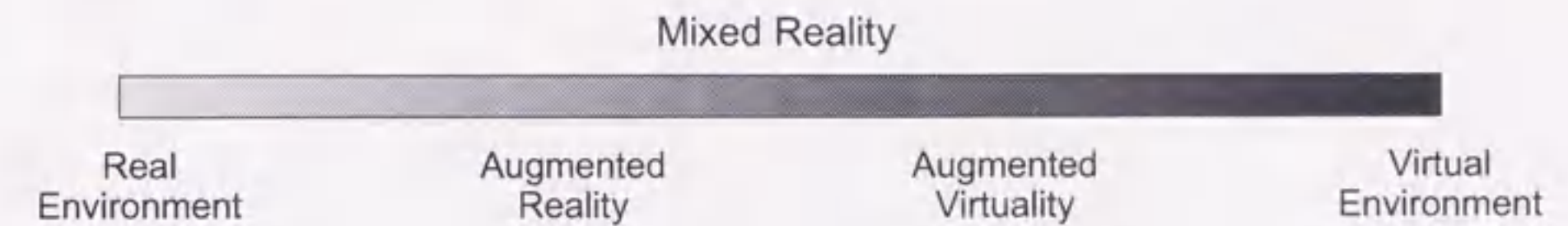


図 1.2: Virtual Continuum.

において最も比重が高い視覚的融合には多くの研究がなされている[14][15]。

## 1.3 Mixed Reality の整合問題

Mixed Reality の基礎的研究は、実空間と仮想空間を如何に高度に融合するかが課題となっている。そのためには、両空間において整合すべき要素がいくつかあり、次のように分類することができる[16][17]。

- 空間的整合

合成した結果そこにあるべきものが正に存在するように、実空間と仮想空間の座標変換を正しく行わなければならない。

- 時間的整合

入出力や伝送の遅延などにより、実空間と仮想空間で時間的ずれが生じる。このずれをできるだけ小さくしなければならない。

- 質的整合

視覚では画質、聴覚では音質といった、仮想空間で生成された感覚が実空間のものと同じ品質を持たなければならない。

この中でも最も基本的な整合問題である空間的整合問題は、特に多くの研究成果が報告されている[18][19][20][21]。



## 1.4 隠蔽関係の表現方法の分類

視覚的融合では、空間的な精度に加えて、物体の相互の位置関係に関する手がかりとして、立体視による距離感とともに隠蔽関係の正確な表現が非常に重要である [22][23].

ここで、実空間と仮想空間とを視覚的に融合することを合成と呼ぶことにして、実空間と仮想空間の隠蔽関係の表現方法を次のように分類することにする.

### i) 単合成

一般的なスーパーインポーズのように、物体の形状や奥行きに関係なく後映像の前に前映像が合成される. 隠蔽関係を固定的に与えておき、テレビのテロップや建物の内部構造の提示 [24] のように片方の映像が他方の映像を視覚的に情報を増強させるときに用いる. 立体視の場合には、隠蔽関係と位置関係の不整合を強く感じることもある.

### ii) ピクセル合成

対象空間中の物体を、視点からの距離の情報 (Depth-Key) として把握することによって仮想空間と実空間の合成する方法である. 視点からの距離を実空間と仮想空間とでピクセルごとに比較して、手前にある空間映像を選択することによって、原理的には正確な隠蔽関係の表現が可能である [25][26].

図 1.3 は、異なる 2 空間の物体 A と物体 B をピクセル合成する例である. それぞれの空間では Depth-Key を持っており、図 1.3(a) のようにピクセルごとに Depth-Key を比較して手前にある方を選択すると、図 1.3(b) 太実線の部分が表示される.

ここで本質的な技術は、距離情報の取得である. 最近ではリアルタイムに近い処理が可能なシステムも開発されているが [27][28], 解像度, 精度, 処理速度はトレードオフの関係にあり, これらを同時に満たすシステムは未だ研究段階である.

## 1.4. 隠蔽関係の表現方法の分類

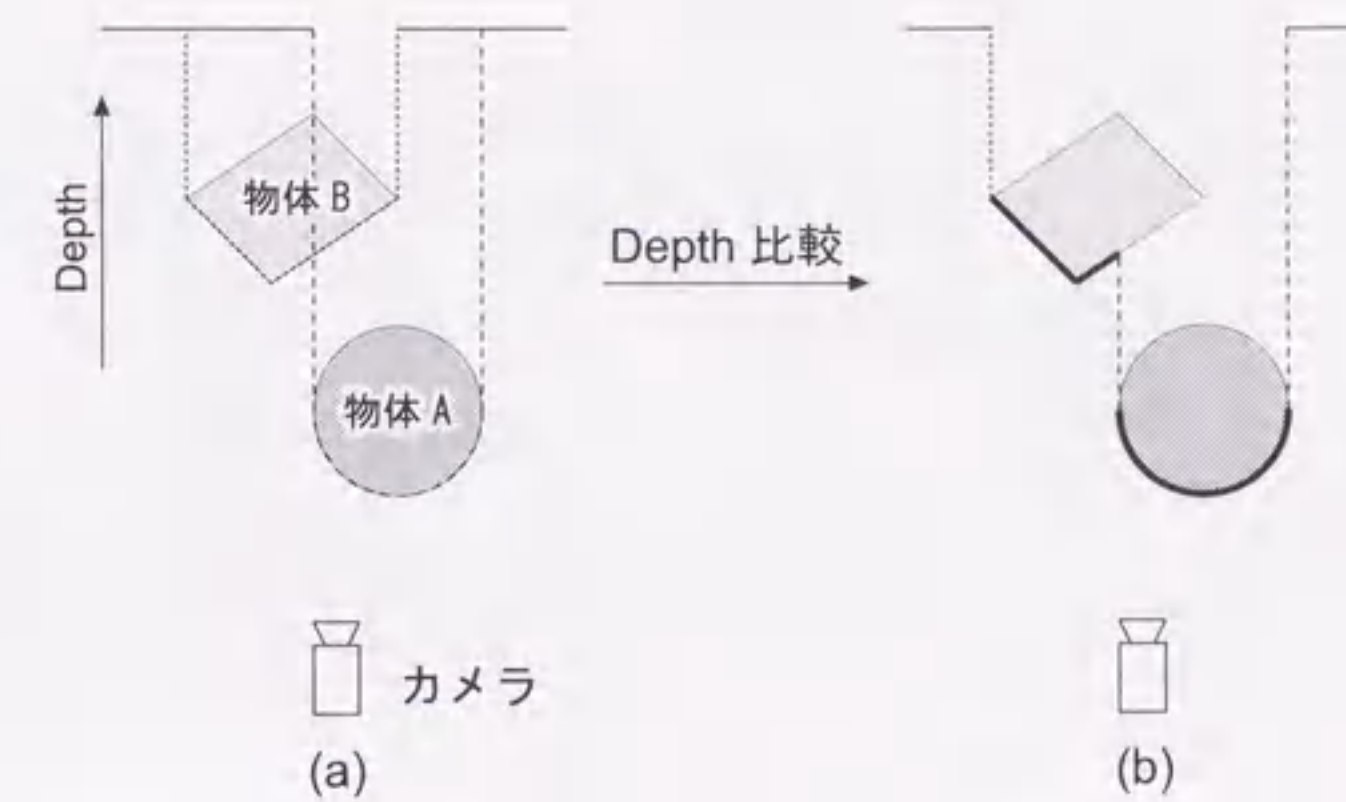


図 1.3: ピクセル合成.

### iii) レイヤー合成

対象空間を複数の層に分割して、それぞれの層の映像をクロマキーなどで合成する. 各層は物体の代表位置に配置され、物体の形状情報は欠落する. このため、物体が離散的に存在すると見せる状態では有効である. 各層の前後関係で隠蔽関係を表現するため、層の存在する物体の位置を把握するのみでよいが、複雑な隠蔽関係を表現することはできない.

図 1.4 は、図 1.3 と同じ条件でレイヤー合成する例である. カメラからの物体 A, B の映像をそれぞれ層 A, B に与えてクロマキー合成すると、図 1.4(b) 右太実線の部分の映像が表示される.

層の映像の合成は平面的であるが、両眼立体視することによって視覚的に三次元合成の効果を得ることができる. また、層の面の決定には既存の高速・高精度な位置検出手法を応用でき、平面映像の合成であるため一般に入手可能なクロマキー合成装置を利用できる.

レイヤー合成を用いた商用バーチャルスタジオ等では、各層にスタジオセットのイメージ (書割) が与えられており、人物とスタジオセットの前後関係に従って人物層と書割層を並び替えることにより三次元合成を実現している [29].



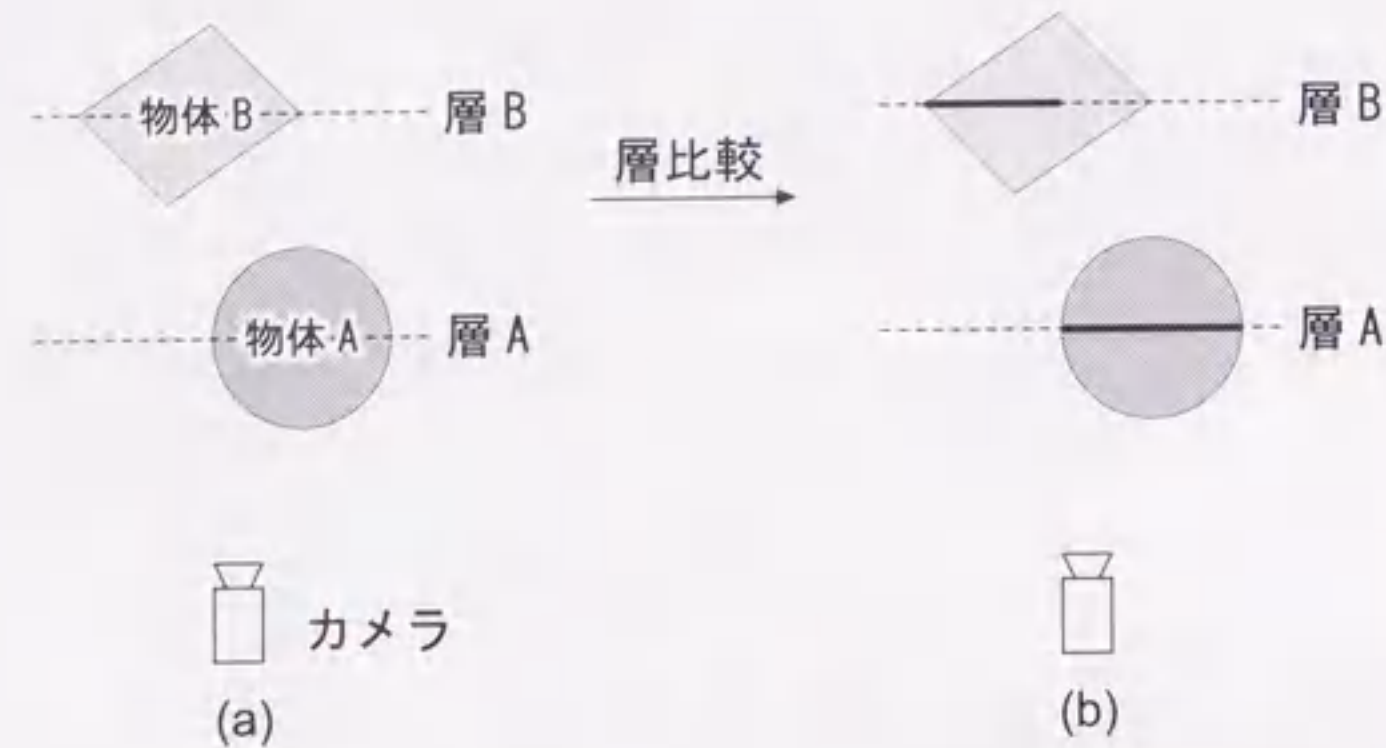


図 1.4: レイヤー合成.

## 1.5 本論文の目的

前節を表現力の面からまとめると、Depth-Keyを使ったピクセル合成は、すべての状況下で正確な隠蔽関係の表現ができる。一方、レイヤー合成は、対象空間に比べて物体が小さく体積を無視できる場合や物体が少ないときなど限られた条件の下であれば、正確に隠蔽関係の表現をすることができる。

また、現在システムの構築の面から見ると、ピクセル合成システムは研究段階であり、合成目的として一般的に入手できる状況に至っていない。一方、レイヤー合成システムは、広く用いられているクロマキー合成装置を使用することができるので、ピクセル合成システムに比べて容易にシステムを構築することができる。

このような理由から、商用合成システムにはレイヤー合成システムが採用されている。しかし、前節で述べたように、現在のレイヤー合成システムは、書割の数だけ層が必要であり、さらに連続的な仮想物体に実物体が囲まれた場面は表現できないという問題点がある。

この問題点を解決するには、ピクセル合成システムを導入すればよいが、商用レベルまで改良されるにはまだ時間がかかると考えられるため、現在使用されているレイヤー合成システムを用いて、表現力を向上させる工夫が必要である。

このような状況から、本研究ではレイヤー合成のシステムの隠蔽関係の表現力を向上さ

せることを目的とする。

まずレイヤー合成システムの原理と表現力について検討する。層の数や層の向きなどと隠蔽関係の表現力の間にはどのような関係にあるか考察することによって、現在のレイヤー合成システムの問題点を明らかにし、改良方法の提案とその効果を述べる。

次に、実際にクロマキー合成装置を用いてレイヤー合成システムを構築し、提案方法をどのように実現するかを説明する。提案方法の実現だけでなく、レイヤー合成システムで重要な位置的整合をとる方法を示す。

最後に、提案方法を導入したレイヤー合成システムを用いた応用例を挙げ、提案方法の効果と有効性を検討する。

## 1.6 本論文の構成

第1章は本章であり、実空間と仮想空間の合成方法と、現在のレイヤー合成方式の問題点について整理し、本論文の目的を述べる。

第2章では、レイヤー合成方式の層による表現力の違いを考察し、既存のレイヤー合成システムを拡張する方法を提案する。そして、提案方法が既存の合成方法に対してどのような位置付けになるかを述べる。

第3章では、提案方法の実現方法について述べる。ここでは、3層のレイヤー合成システムを構築し、そのシステムでの実空間と仮想空間の座標系の整合方法について説明する。

第4章では、構築したレイヤー合成装置を用いて、二つの応用例を挙げて提案方法の有効性を示す。

最後に第5章では本論文での成果をまとめている。



## 第2章 多層分割による隠蔽関係の表現

レイヤー合成の基本的な手順は、対象空間をいくつかの層に分割することと、それらの層を合成することである。本章では、レイヤー合成の基本原則と表現力についてまとめ、より表現力を高めるための方法を提案する。

### 2.1 レイヤー合成の原理

原理を説明するために、例として車・人・ビルの3つの物体が存在し、これらの物体を並べて図2.1のような映像を生成する場合を考える。



図 2.1: 対象空間の例.

このような映像を生成する場合には、次のような手順を行う。

- (1) 映像中の物体の前後関係から、図2.2のように映像空間を層 A・B・C の3層に分割する。
- (2) 図2.3のような各層の映像をカメラ撮影やCGなどで生成する。
- (3) 図2.4のように層の奥行き順に合成する。



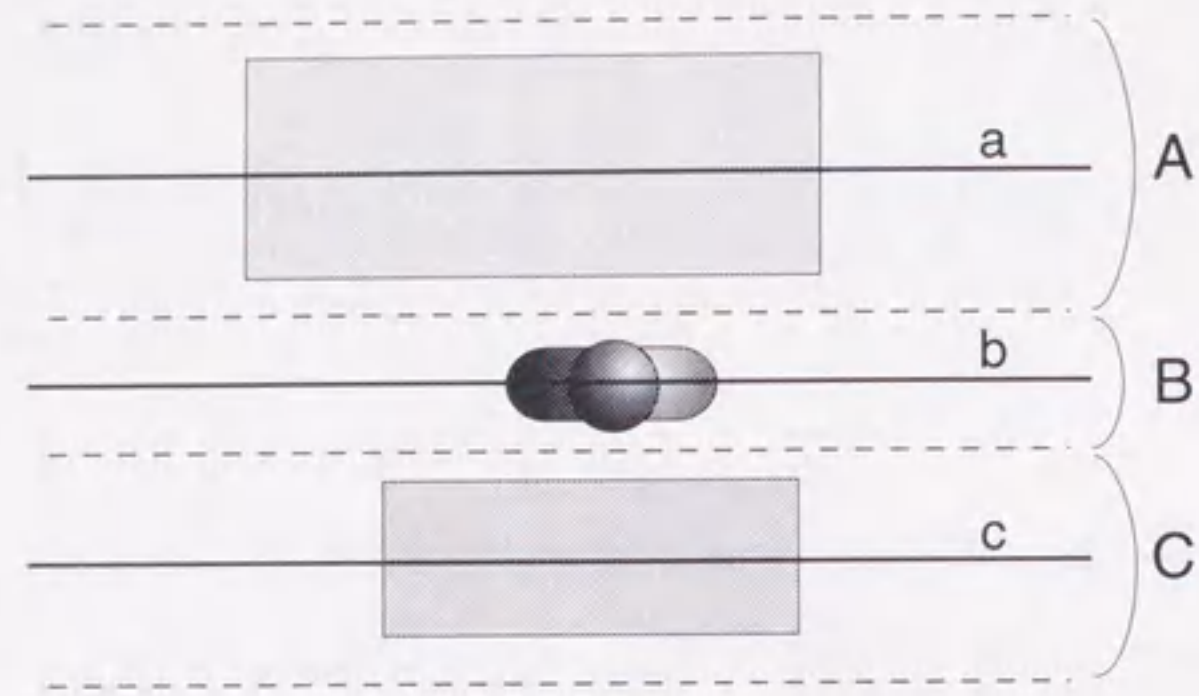


図 2.2: 対象空間の分割.

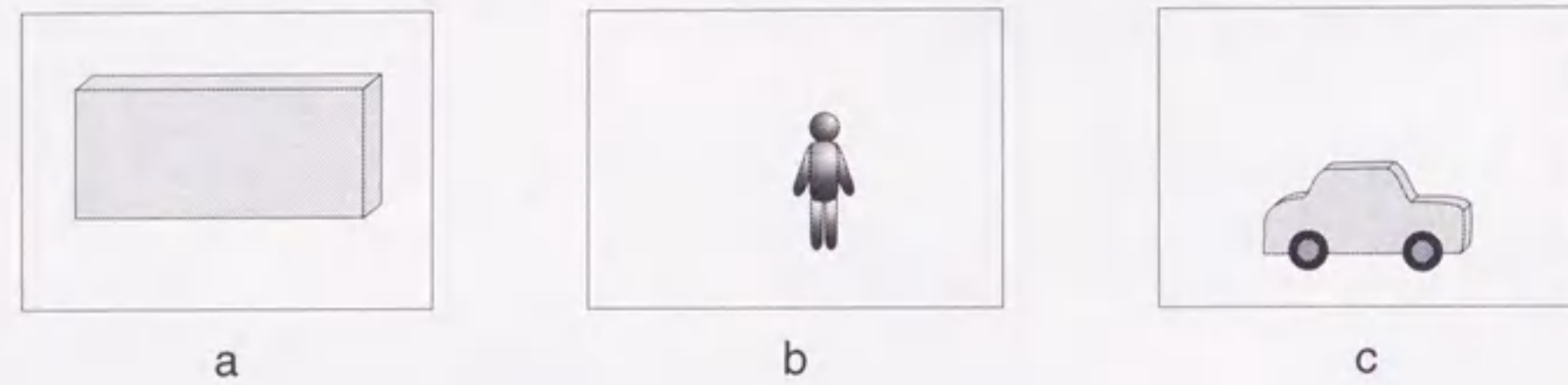


図 2.3: 各層の映像.

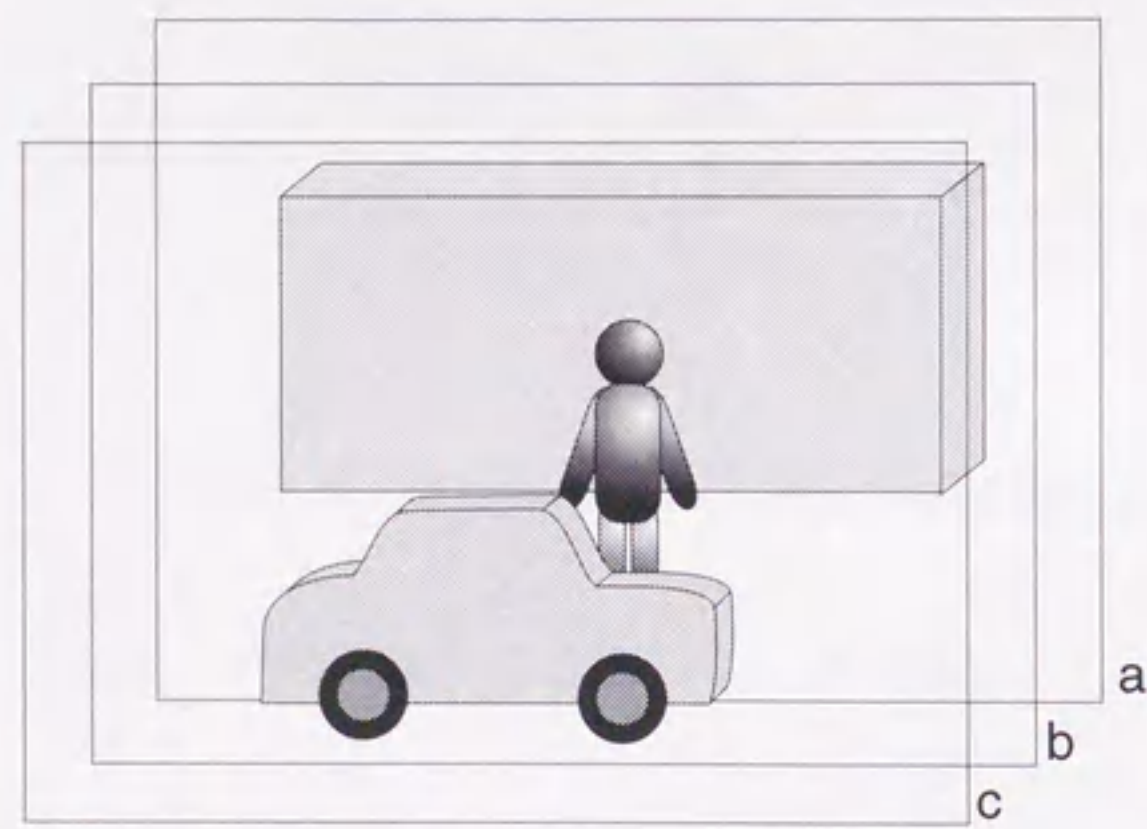


図 2.4: 分割空間の合成.

ここで、車とビルの層をCGで作成し人をカメラで撮影すれば、人が2つのCGで作られた空間の間に存在しているような効果を得られる。

## 2.2 レイヤー合成の表現力

本節では、既存のレイヤー合成システムを発展させるために、レイヤー合成の表現力が層の数・向きなどによってどのような違いがあるかを考察する。

### 2.2.1 垂直分割

垂直分割とは、視線方向に対して垂直に層を分割した場合である。一般にレイヤー合成をするときは、対象空間を垂直分割していると考えることができる。

#### i) 垂直2層分割

図2.5は、対象空間が実空間と仮想空間の2つの層からなる場合で、一般的なクロマキー合成装置を用いたときの層である。この場合では手前と背後という関係を表現することができる。

第1.4節で述べたように、Augmented Realityでは、実空間を補う情報をCGとして仮想空間を作成し、実空間映像の前に合成して実空間の情報量を増すときに用いられることがある。

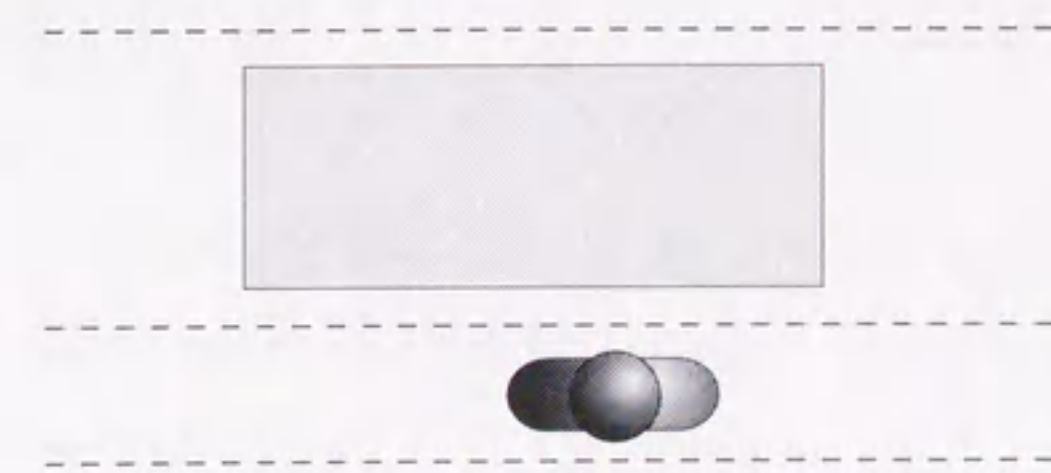


図 2.5: 垂直2層分割



## ii) 垂直3層分割

第2.1節の説明で用いた図2.1は垂直3分割の例である。層Bが実空間で層A, Cが仮想空間の場合では、仮想空間の物体の間に実空間の物体が入り込んでいるかのように見える。また、図2.6のように層A, Cが連続的な場合でも、適切に層を決定することによって正しく合成することができる。

しかし、逆に層A, Cが実空間で層Bが仮想空間の場合には、物理的に層Aと層Cを分離して層の映像を生成することができないため、表現不可能である。実空間の物体が層を跨いで連続でなくても、一般に実空間を同一の視点から撮影して異なる層に分割することは困難である。そのため、実空間を複数の層とする時は、遠隔実空間の協調作業など、実物体は層を跨ぐことのない異なる実空間の物体を考えると不具合が生じにくい。

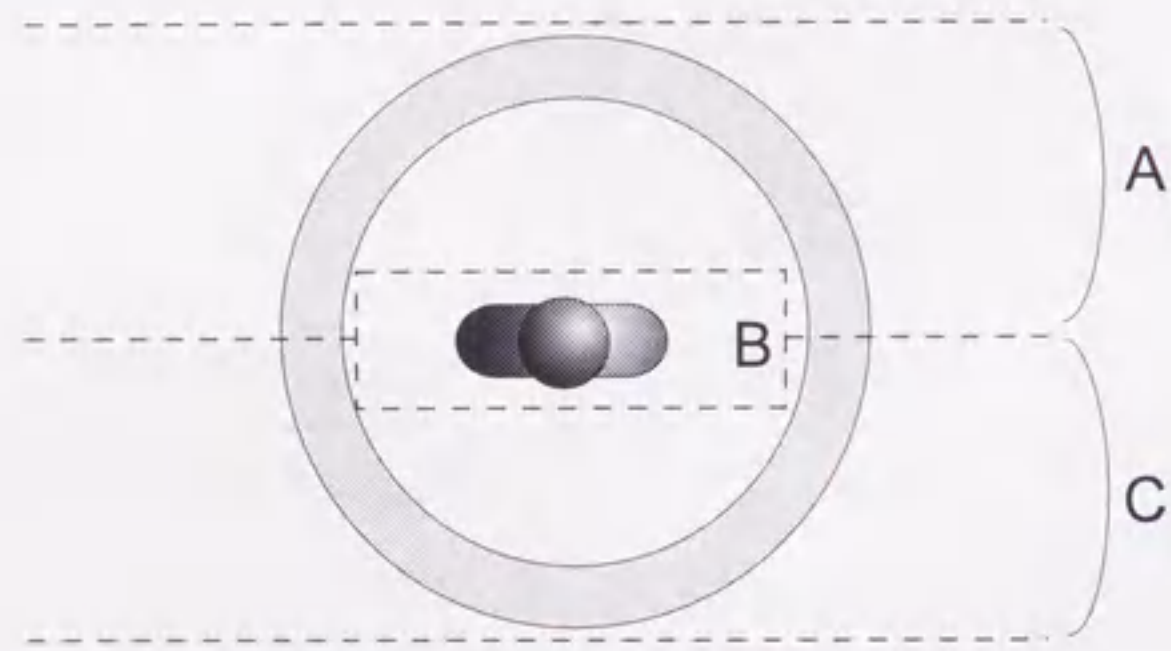


図 2.6: 連続体の垂直3層分割

iii) 垂直  $n$  層分割

対象空間を  $n(n = 2m \text{ or } 2m + 1, m \in \mathbb{N})$  層に分割した時、実空間が  $m$  層、仮想空間が  $m$  層もしくは  $m + 1$  層とすると、仮想空間空間に  $m$  個の実空間の物体が入り込んで隠蔽関係を表現することができる。

## 2.2.2 層の向きによる違い

書割合成システムは垂直分割の層で構成されており、合成する物体の数だけ層が必要である。ここで、層の向きを視線に対して垂直に固定ではなく、斜めに決定することを考える。

## i) 斜面2層分割

図2.7のように対象空間を斜めに2分割した場合、垂直2層分割と同じ様に実空間と仮想空間の前後関係のみを表現することができる。

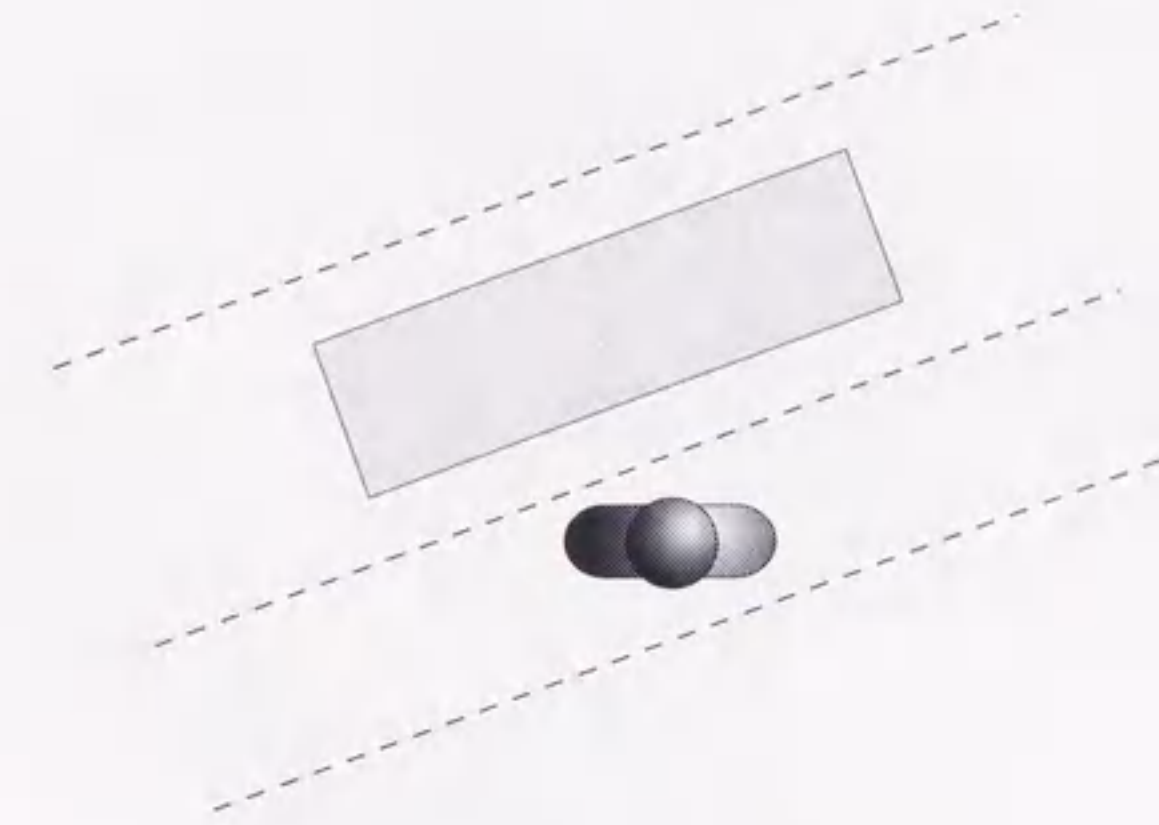


図 2.7: 斜面2層分割.

## ii) 斜面3層分割

垂直3層で述べたように、3層分割では実空間が1層、仮想空間が2層の場合が現実的であるため、斜面3層でも同じ層の構成を考える。

まず、Fig.2.8のように単純に斜めに3層に分割した場合は、垂直3層分割と同じ効果を得ることができる。

ここで、実空間の層にある物体が対象空間に対して十分に小さい場合、図2.9のように、2つの物体を含む層で分割することによって、3層分割でありながら垂直5層分割に相当す



る効果を得ることができる。さらに、3つの物体に対してそれらを含む層で分割すれば、垂直7層分割に相当する効果を得ることができる。

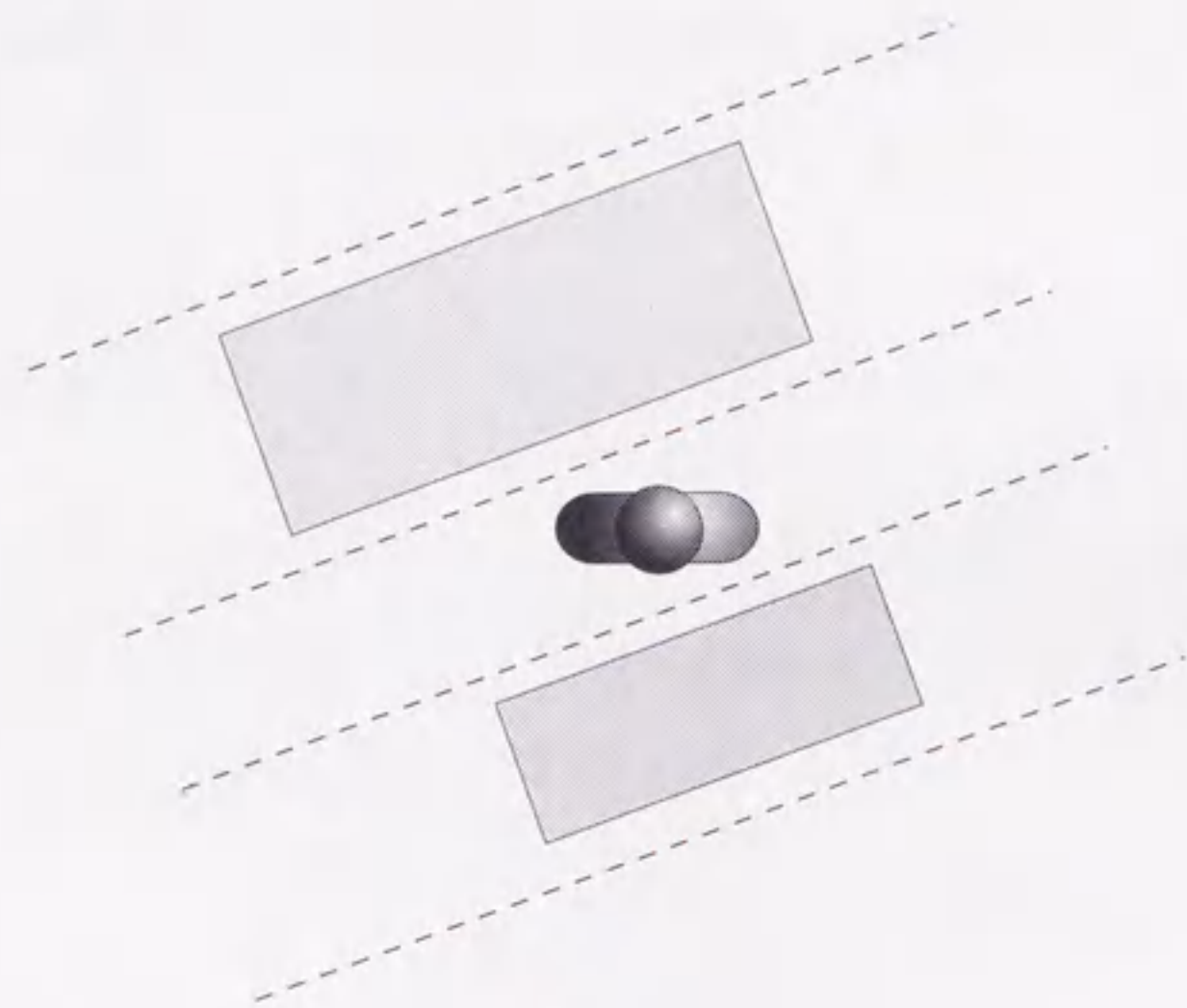


図 2.8: 斜面3層分割.

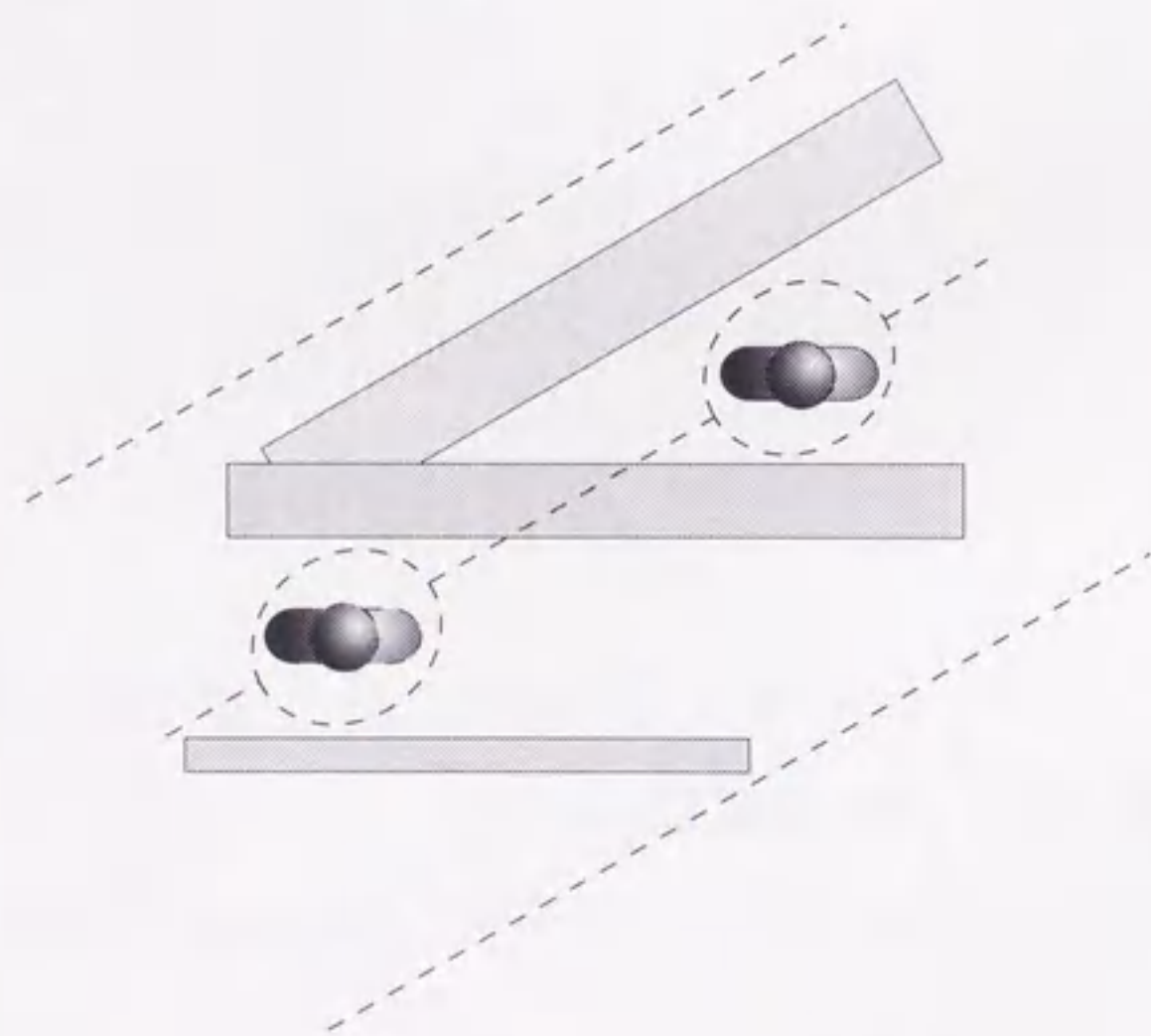


図 2.9: 斜面3層分割による垂直5層分割表現.

すなわち、体積が十分小さい  $n$  個の物体が同一平面上にあるとすると、垂直  $2n+1$  層分割と同じ効果があることになる。そのため、空間の構成しだいでは層の削減に大きな効果がある。ここで  $n \rightarrow \infty$  とすると、平面を持つ実物体となる。したがって、仮想空間に板などの平面型の物体を挿入することもできることが分かる。

### 2.2.3 静的な層・動的な層の違い

図 2.10 のような、実空間1層・仮想空間2層の場合を考える。静的に層を決定すると、その層に登場する物体（人）は図 2.10 の斜線範囲の領域しか移動することができない。

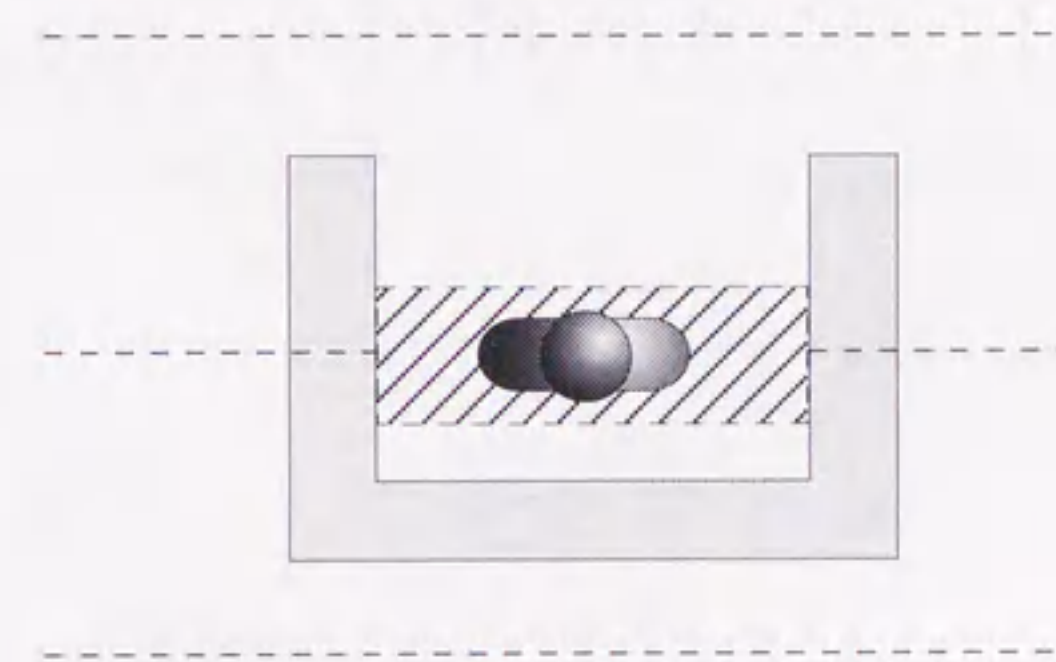


図 2.10: 静的な層.

ここで、層を動的に決定することを考える。いま、図 2.11 のように、人の位置に合わせて層を  $1 \rightarrow 2 \rightarrow 3$  と変えることで、人は囲いから外へ出ることができる。

このように、層を動的に決定することは物体の移動範囲の拡大に有効であると考えられる。

### 2.2.4 単眼視・立体視の違い

例えば図 2.12 のように、層 A の物体（ビル）と層 B の物体（人物）の映像がそれぞれ合成面 a, b にある場合を考える。

ここで、図 2.13 のようにビルと人が干渉したとする。単眼視の時には合成映像に何ら違



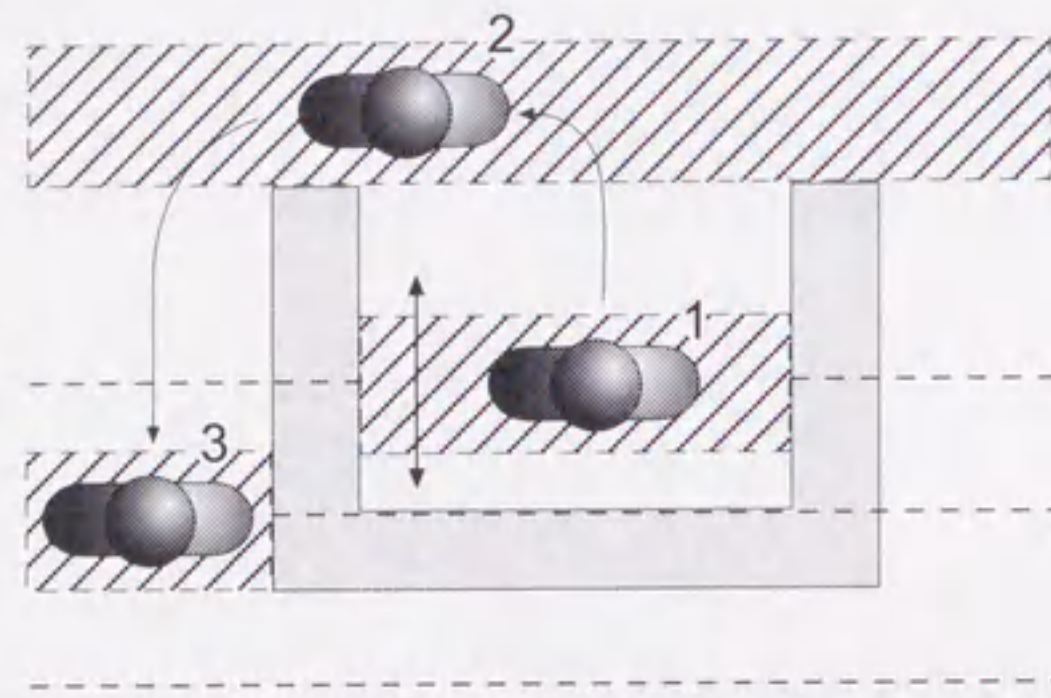


図 2.11: 動的な層.

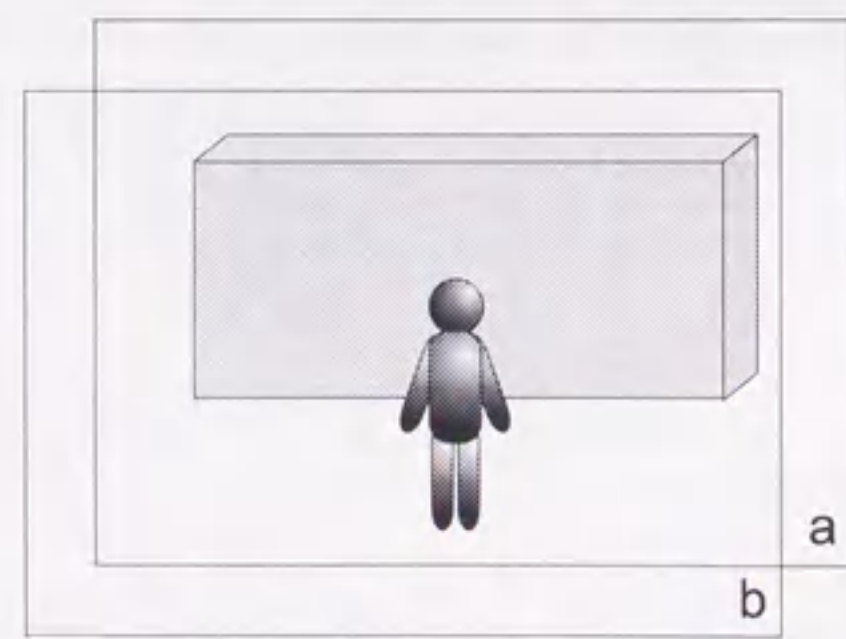
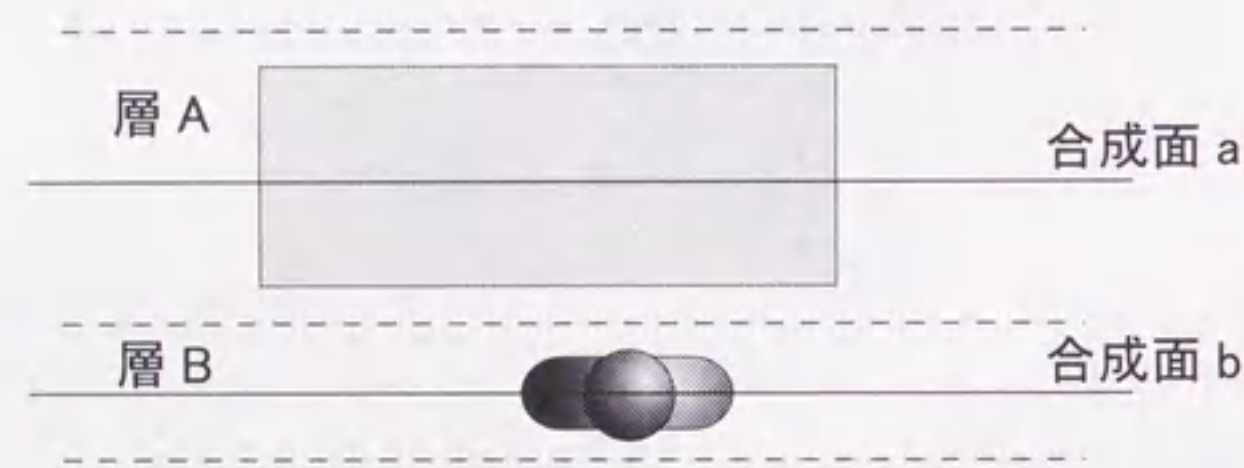


図 2.12: 正常な隠蔽関係.

和感はない。なぜならば単眼視の奥行き感は、物体の隠蔽関係や幾何学的理解によるところが大きいからである。一方、立体視では単眼視の奥行き感に加えて、物体のボリューム感を感じるので、人とビルの前後関係に違和感を感じうる。

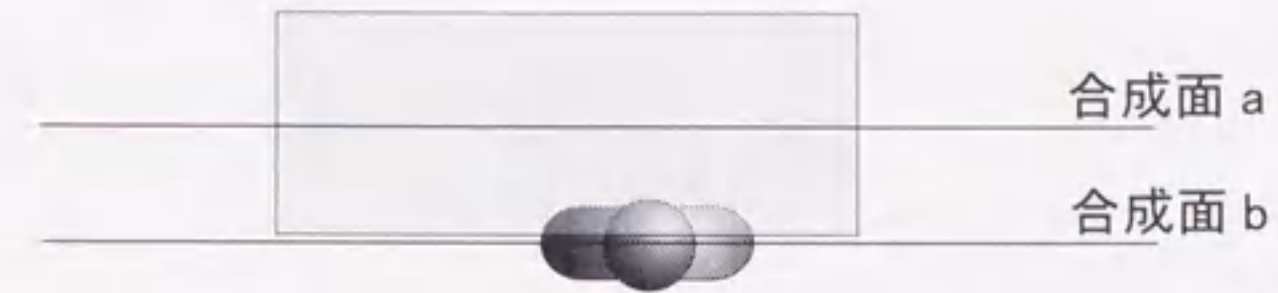


図 2.13: 物体同士の干渉.

したがって、単眼視の時には層の平面的な映像の前後関係による隠蔽関係の表現であるが、立体視の時には平面的な映像ではなく、もともと体積を持った物体同士の三次元的な隠蔽関係の表現といえる。このため、レイヤー合成でありながら、ピクセル合成に近い表現力を得ることができる。

## 2.3 改良方法の提案

### 2.3.1 従来のレイヤー合成システムと改良点

この節では、前節の考察から従来のレイヤー合成システムの改良点を述べる。

商用システムも含めてこれまでのレイヤー合成システムは、垂直分割のレイヤー合成システムである。特に、商用システムの市場はバーチャルスタジオ用であるため、カメラワーク・映像品質・作成コストに重点をおいており、垂直分割型で目的を果たしている。また、人物とCGで作成されたバーチャルセットの前後関係が入れ替わる時は、人物とバーチャルセットの層を入れ替えることで隠蔽関係の整合を保っている。

そこで、斜面分割・動的な層の決定・立体視を導入することによって

- (1) 必要な層の削減
- (2) 連続した仮想物体への実物体の挿入



- (3) 移動範囲の拡大
- (4) 視覚的な三次元融合

の効果を得ることができると考えられる。本論文では、このような点から、既存のレイヤー合成システムの隠蔽関係の表現力を向上させる方法を提案する。

### 2.3.2 提案方法の位置づけ

ここで、前節の考察から本論文で提案する方法が他の合成方法と比較してどのような位置付けになるか考えると、距離情報空間において図 2.14 に示すように分類できる。(a) は実物体と仮想物体の両方の距離情報を求めて Depth-Key による合成を行っている。しかし先に述べたように、実物体の距離情報の取得手段は研究段階である。そのため、実物体は代表距離にあるものとみなすことによって Depth-Key による合成を簡略にしたのが (b) である。実物体が対象空間に対して体積を無視できるときは (a) と同じ効果があるため、商用システムでも応用された例がある。(d) は仮想物体映像と実物体映像をレイヤー合成する方法で、実物体・仮想物体は共に体積がなく、実物体が仮想物体で囲まれた隠蔽表現はできない。本提案方法は (c) にあたり、仮想物体の体積を考慮し、仮想物体映像は実物体映像の層の位置によって動的に決定される。(d) と同じようにレイヤー合成であるが、(b) と同じ条件下で (a) と同等な隠蔽表現を得ることができる。

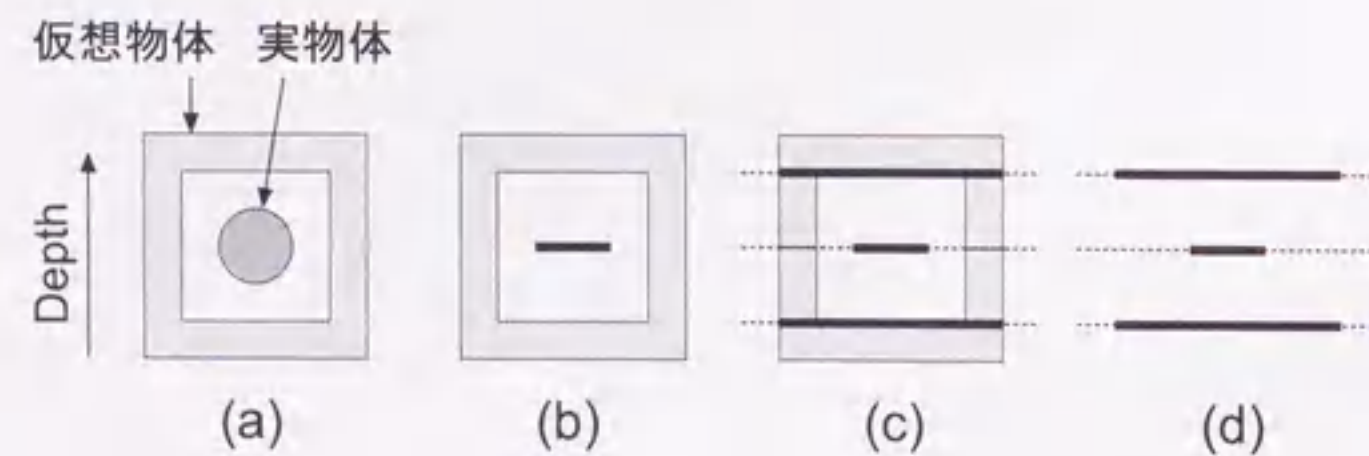


図 2.14: ピクセル合成とレイヤー合成の分類。

## 第3章 3層分割を用いた実空間と仮想空間との合成方法

本章では、まず、前章で述べた3層分割を用いて実空間と仮想空間との合成の実現するための方法について提案する。そしてレイヤー合成システムの構築と、システムの効率化・高精度化について述べる。

なお、本実験では3層は実空間1層、仮想空間2層とし、仮想空間の中に実空間が挿入される場面について考える。

### 3.1 合成方法

ここでは、図 3.1 に示すように仮想空間 (CG) に実空間 (人物) が入り込んだ状態を表現することについて考える。

まず、人物をカメラで撮影し、コンピュータで前景 CG と後景 CG を生成することによって、それぞれの層に割り当てられる映像を準備する。前景 CG と後景 CG は、仮想空間を人物の位置を境界面 (クリッププレーン) とし前後に切り分けて生成される。

次に、2つの CG 層の間に人物層が挿入された3層を構成するために、まずカメラ映像を後景 CG の前に合成し、さらにその合成映像の前に前景 CG を合成する。この結果、CG 空間の中に人物が入り込んでいるような効果を得ることができる。

さらに、左右両眼用に2系統の映像についてこのような処理をすることによって、立体視可能なステレオ映像を生成することができる。このとき、隠蔽関係は平面3層であるが、立体視をすることにより、3次元の空間同士が融合しているような効果を得ることができる。



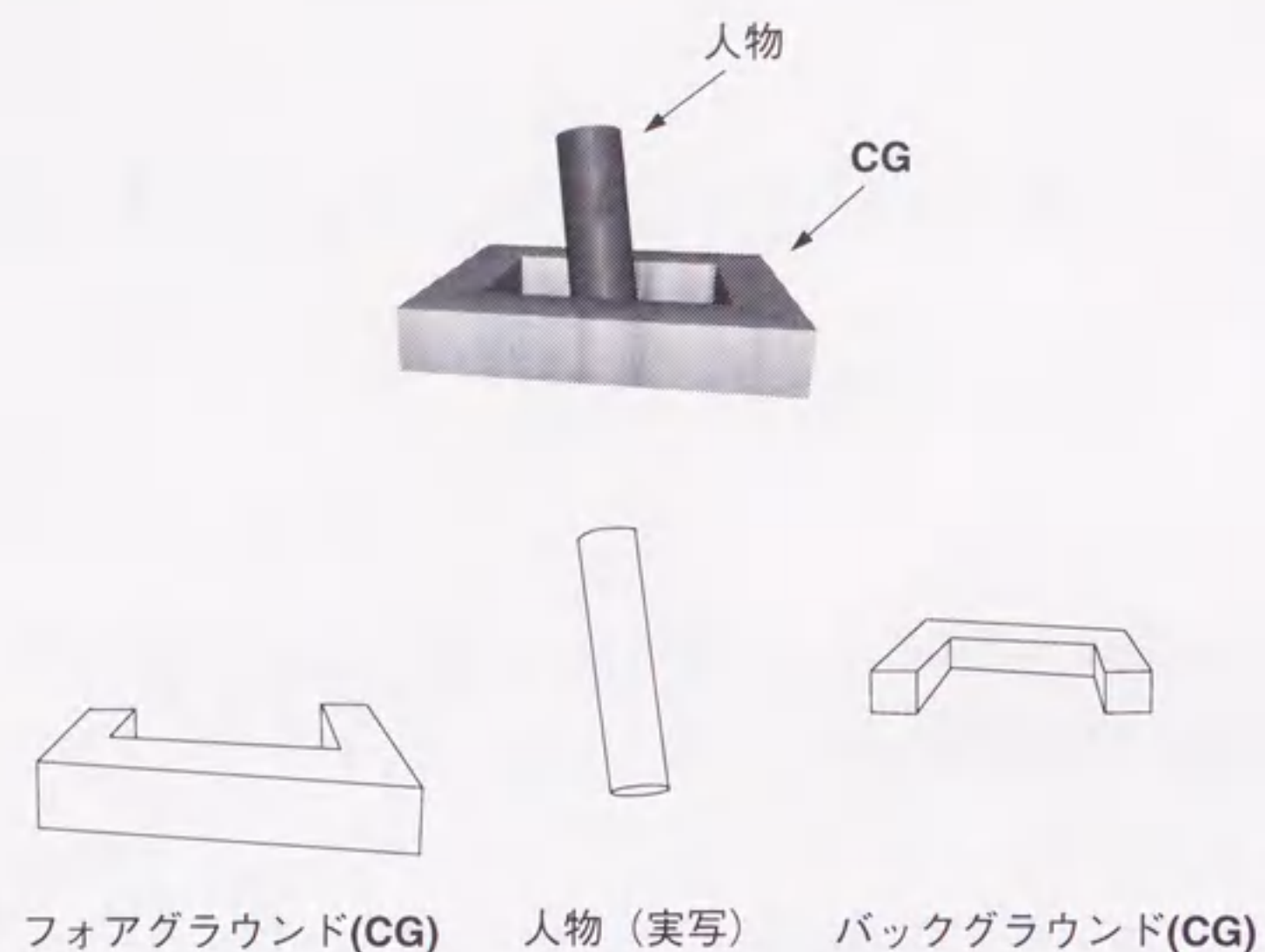


図 3.1: 3層の生成.

### 3.2 クリッププレーンの決定

正しい隠蔽関係を得るためには、仮想空間を2つに切り分けるクリッププレーンの決定方法が非常に重要である。

#### i) 実物体が一つで体積を無視できるとき

対象空間に対して実物体(人物)が十分小さく体積を無視できるときは、図3.2のようにカメラに対して垂直な面をクリッププレーンとする。このとき実物体の位置はカメラの光軸方向(深さ方向)の座標検出のみでよいので、簡単に高速な処理を期待できる。

#### ii) 実物体が一つで体積を無視できないとき

しかし、後述する図4.11のように手や道具を用いた組立作業をこの対象空間で行うような、実物体が対象空間に対して十分小さいと見なせない場合は、前節で述べたような実物

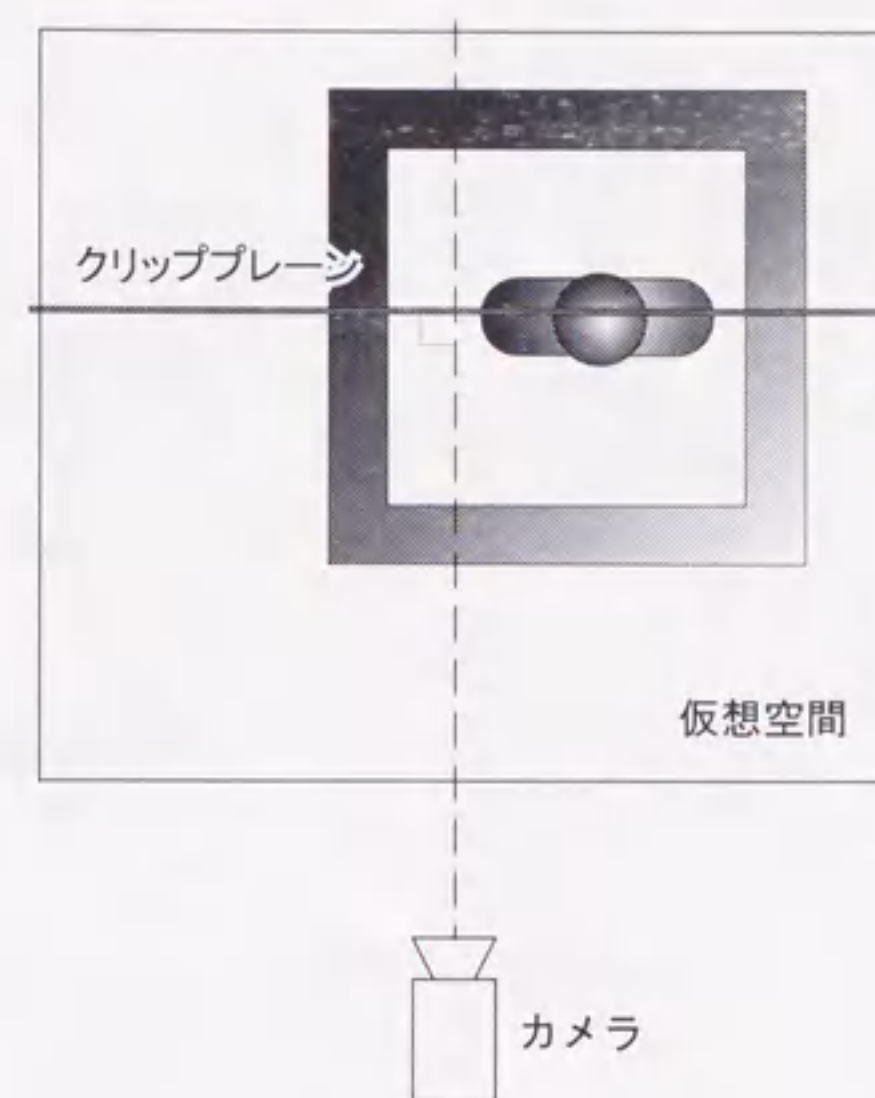


図 3.2: 一人の場合.

体の代表位置にカメラに対して垂直な面をクリッププレーンと隠蔽関係の不整合が顕著となることがある。

図3.3の例では、Plane 1のようにカメラに対して垂直な面を用いると、空間での実際の配置と合成結果から得られる隠蔽関係に矛盾が生じている。そこで、この例ではPlane 2のように、実物体(この場合は手や道具)に合わせてクリッププレーンを決定すると正しい隠蔽関係を持った合成結果を得ることができる。

#### iii) 実物体が二つで体積を無視できるとき

図3.4のように人物が二人の場合は、二人の間にカメラに対して斜めにクリッププレーンを決定することによって、層を増やすことなく仮想空間に入り込むことができる。

このようにクリッププレーンの取り方を工夫することで、少ない合成回数でより複雑な隠蔽関係の表現力を得ることが可能である。



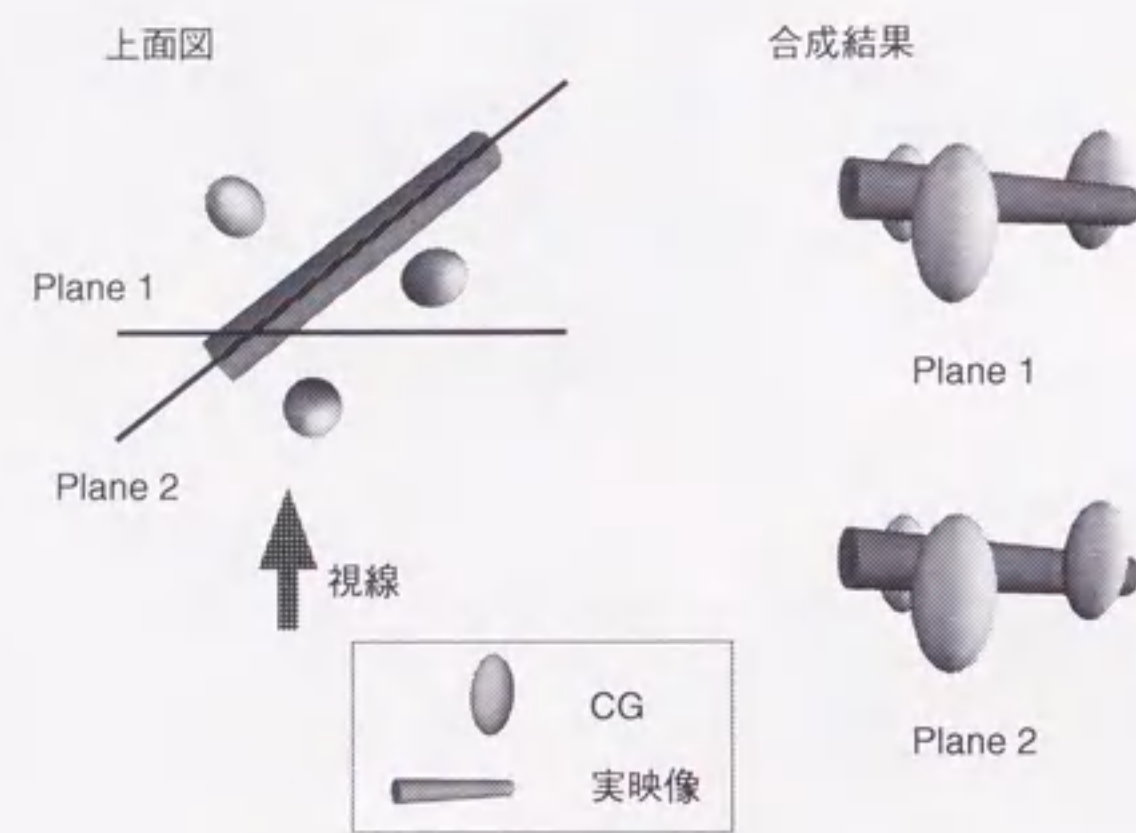


図 3.3: クリッププレーン.

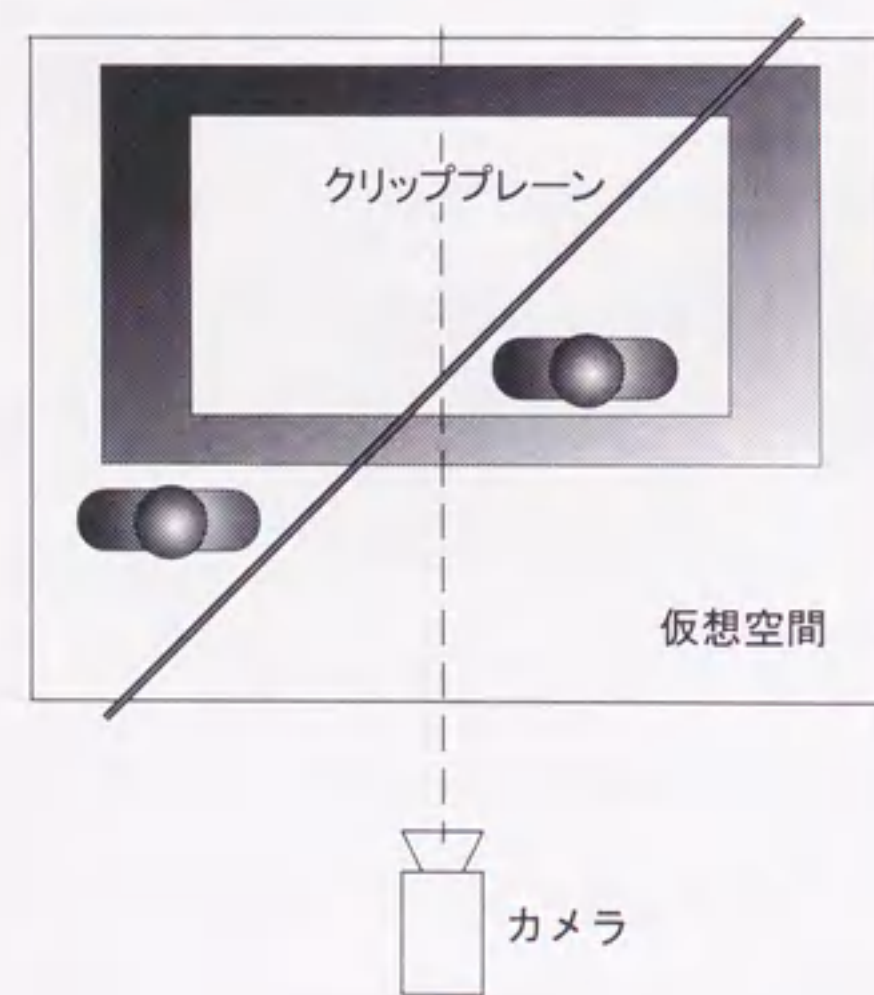


図 3.4: 二人の場合.

### 3.3 システム構成

本研究では、岐阜大学バーチャルシステム・ラボラトリー内のバーチャルスタジオに設置されているVGA 双眼三画面クロマキー合成システムを使用した[30][31]. このシステムはすべて映像をVGAで扱っており、CGとの親和性がよい. 図3.5はシステム構成の概要で、大きく次のように分類される. なお、本システムを構成する機器の主な仕様を表3.1～表3.5に示す.

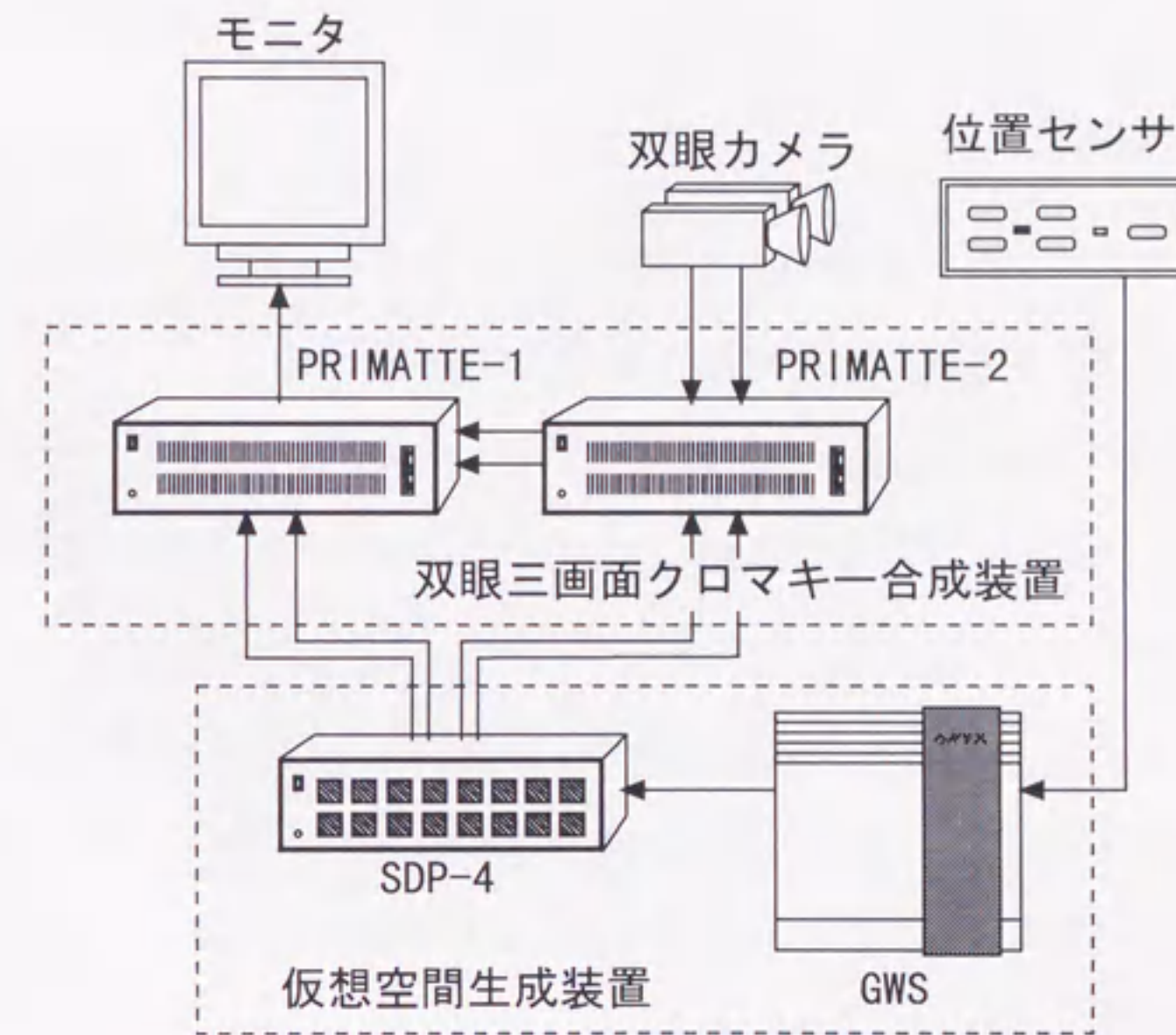


図 3.5: システム構成.

#### i) 双眼三画面クロマキー合成装置

双眼三画面クロマキー合成装置は、PRIMATTE-1(表3.1)とPRIMATTE-2(表3.2)を直列に接続して構成されている. PRIMATTE-1はVGA 2系統4入力(前景用2入力, 後景用2入力), 倍速VGA 1出力(フィールドシーケンシャル立体合成映像)を持つ. PRIMATTE-2はVGA2系統4入力(前景用2入力, 後景用2入力), VGA 2出力(左右合成映像)を持つ.



PRIMATTE-2で Graphics Work Station (GWS) の後景CGとVGAカメラの入力を合成し、その合成映像とGWSの前景CGをPRIMATTE-1で合成し、倍速VGAでモニタに出力する。

PRIMATTEのOSはmicrosoft社windows95を採用しており、制御プログラムには、PRIMATTE PRO100 GUI for Windows95を用いている。このプログラムはGUI(Graphical User Interface)によりクロマキー色(自動・詳細)、前後景切替、出力映像選択などを行うことができる。

#### ii) 仮想空間映像生成装置

この装置は、物理現象などを計算して仮想空間を構成し、前景・後景それぞれの左右両眼用の4映像を生成する装置で、GWSとSDP-4から成る。

GWSにはOnyx RealityStation(表3.3)を用い、グラフィックライブラリにはIrisGLを使用している。このGWSの解像度は1280×1024であるので、GWSの1画面にクロマキー合成装置の入力640×480VGAを4映像表示することができる(図3.6)。

SDP-4(表3.4)は入力画像の任意の4領域を切り出して出力する機能を持っている。切り出す領域や出力解像度などは、他のコンピュータよりRS-232Cを介して設定することができる。SDP-4はGWSで生成された映像をVGA4映像に切り分けて、双眼三画面クロマキー合成装置へ出力している。

#### iii) 実空間撮影用双眼カメラ

これは、仮想空間に入れ込む実空間の映像を立体撮影するためのカメラで、双眼カメラにはVGA60(表3.5)を使用している。VGA60はプログレッシブスキャンのVGA出力を持ち、ズーム・フォーカスなどをRS-232C経由でも操作できる。さらに、VGAタイミングの24bitデジタル出力を持っている。

なお、カメラ映像にもクロマキーが必要になるので、本実験ではブルーバック撮影している。

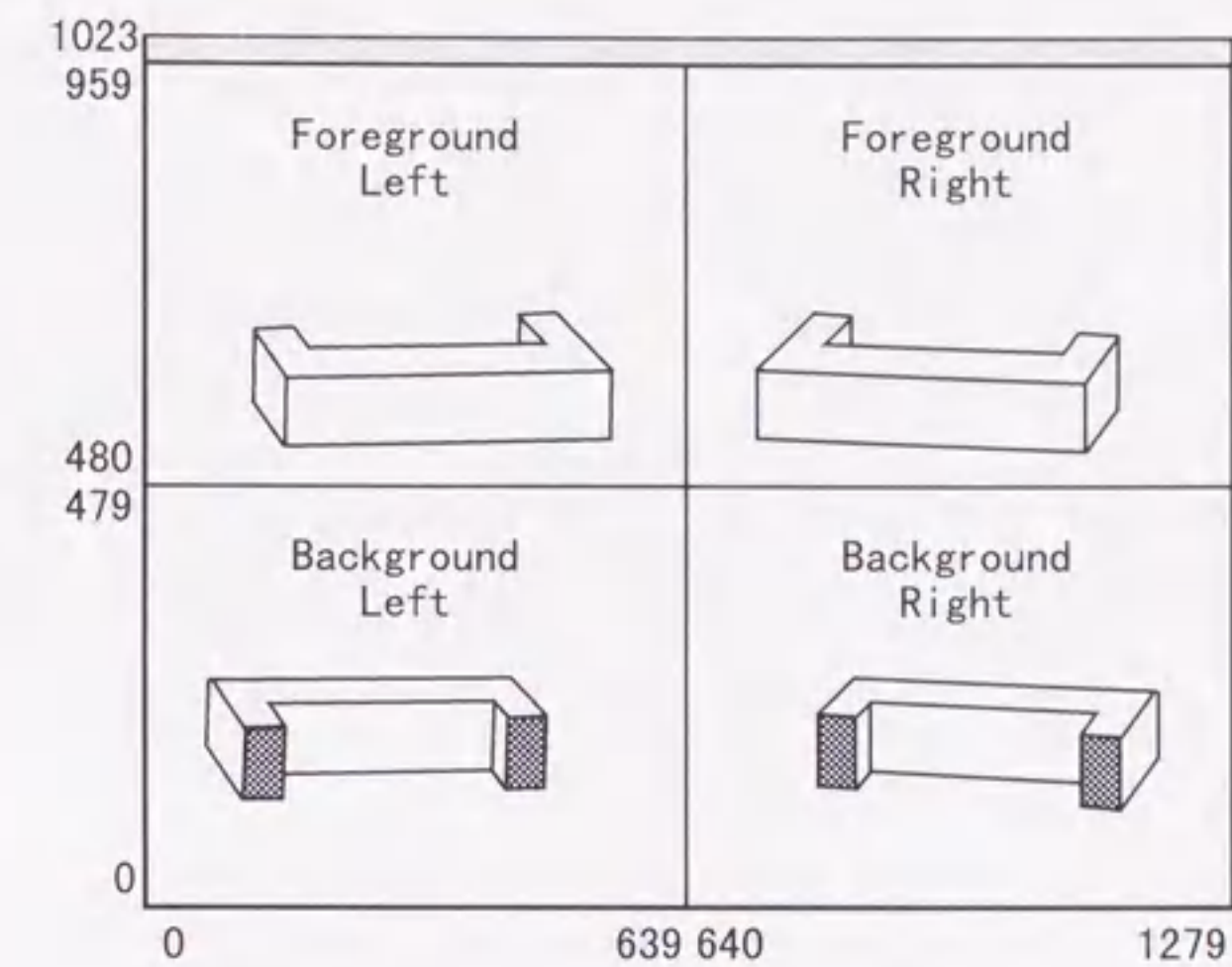


図 3.6: SDP-4用のGWS生成画面。

#### iv) 位置センサ

仮想空間のCGを前景CGと後景CGに切り分けるためのクリッププレーンを求めるために、カメラ映像の被写体の位置を検出しなければならない。本実験では、上部カメラによる位置検出と磁気位置センサによる検出を行った。位置センサについては第4章にて述べる。



表 3.1: PRIMATTE-1 の仕様

項目	仕様
メーカー	株式会社 PHOTRON
入出力	4 入力 1 出力
入力信号	VGA(水平 31.5kHz, 垂直 60Hz, ノンインターレース)
出力信号	2 倍速 VGA(水平 63.0kHz, 垂直 120Hz, ノンインターレース)

表 3.2: PRIMATTE-2 の仕様

項目	仕様
メーカー	株式会社 PHOTRON
入出力	4 入力 2 出力
入出力信号	VGA(水平 31.5kHz, 垂直 60Hz, ノンインターレース)

表 3.3: Onyx の仕様

項目	仕様
メーカー	Silicon Graphics, Inc.
CPU	MIPS R10000 194MHz
FPU	MIPS R10010
メモリ	128MB
グラフィックボード	RealityEngine2 Graphics
OS	IRIX 6.2

表 3.4: SDP-4 の仕様

項目	仕様
メーカー	株式会社 PHOTRON
入出力	1 入力 4 出力
入力解像度	1280×1024 (最大)
入力水平周波数	30kHz ~ 100kHz (NTSC 時 15.75kHz)
サンプリング周波数	25 ~ 130MHz
出力解像度	640×480 (最大)
出力周波数	水平 31.436kHz, 垂直 59.877Hz, ノンインターレース (NTSC 時 水平 15.734kHz, 垂直 59.940Hz, インターレース)
出力ドットクロック	25.45MHz (NTSC 時 12.288MHz)

表 3.5: VGA60 の仕様

項目	仕様
メーカー	株式会社 PHOTRON
CCD	35 万画素 6mm 全画素インターライン転送 CCD エリアイメージセンサ
レンズ	$f = 6.0mm \sim 60.0mm, F1.6 \sim F2.7$
映像出力	VGA(水平 31.5kHz, 垂直 60Hz), アナログ/デジタル



### 3.4 実空間と仮想空間の座標系の整合

実空間と仮想空間を正しく合成するには、前章で述べた方法のほか、両空間の座標系が一致しなければならない(空間座標系の整合性)。そのためには、

- (1) 実空間と仮想空間の座標系の一致
- (2) カメラの姿勢の検出
- (3) カメラとCGの透視変換の一致
- (4) クロマキー合成装置の実映像とCGとの合成位置の補正

のための作業・計算をする必要がある。

これらの整合に必要な変換をまとめて変換パラメータと呼ぶことにする。これらの変換は $4 \times 4$ 行列の回転・平行移動・拡大縮小・透視変換を用いて表現することができる。

クロマキー合成時の実空間と仮想空間の一致とは、結局、実空間(グローバル座標系)と仮想空間で表示装置のスクリーン座標系への変換が一致しているということであるから、変換パラメータのキャリブレーションとは、この変換行列を求めることである。

#### 3.4.1 変換行列の算出

先に述べたように、非線形要素(レンズの光学歪みなど)がないとすると、変換パラメータは $4 \times 4$ 行列で表すことができる。いま、グローバル座標系の点 $(x_g, y_g, z_g)$ 、表示装置のスクリーン座標系の点 $(x_s, y_s)$ 、変換行列を $T$ とすると、

$$\begin{bmatrix} x_g & y_g & z_g & 1 \end{bmatrix} T = \begin{bmatrix} x_s^* & y_s^* & 0 & h \\ x_s & y_s & 0 & 1 \end{bmatrix} \quad (3.1)$$

$$T = \begin{bmatrix} T_{11} & T_{12} & 0 & T_{14} \\ T_{21} & T_{22} & 0 & T_{24} \\ T_{31} & T_{32} & 0 & T_{34} \\ T_{41} & T_{42} & 0 & T_{44} \end{bmatrix} \quad (3.2)$$

である。キャリブレーションでは、既知の $(x_g, y_g, z_g)$ に対応する $(x_s, y_s)$ を示すことで、 $T$ の要素を求める。このとき必要な点は6点である[32]。非同次行列とするために $T_{44} = 1$ とすると条件余剰となるので、行列 $T$ の各要素を最小二乗法で求めている。さらに多くの点でキャリブレーションをすれば、より精度のよい変換行列を求めることができる。

#### 3.4.2 GLパラメータとの関係

IrisGLでは、仮想空間の三次元座標から画面の二次元座標に変換するために、図3.7に示すような座標変換を行っている[33]。IrisGLを用いて仮想空間と実空間の座標系を一致させて合成表示するためには、それぞれの変換を求めなければならない。ここで、IrisGLの座標系と座標変換について説明し、本実験との関係について述べる。

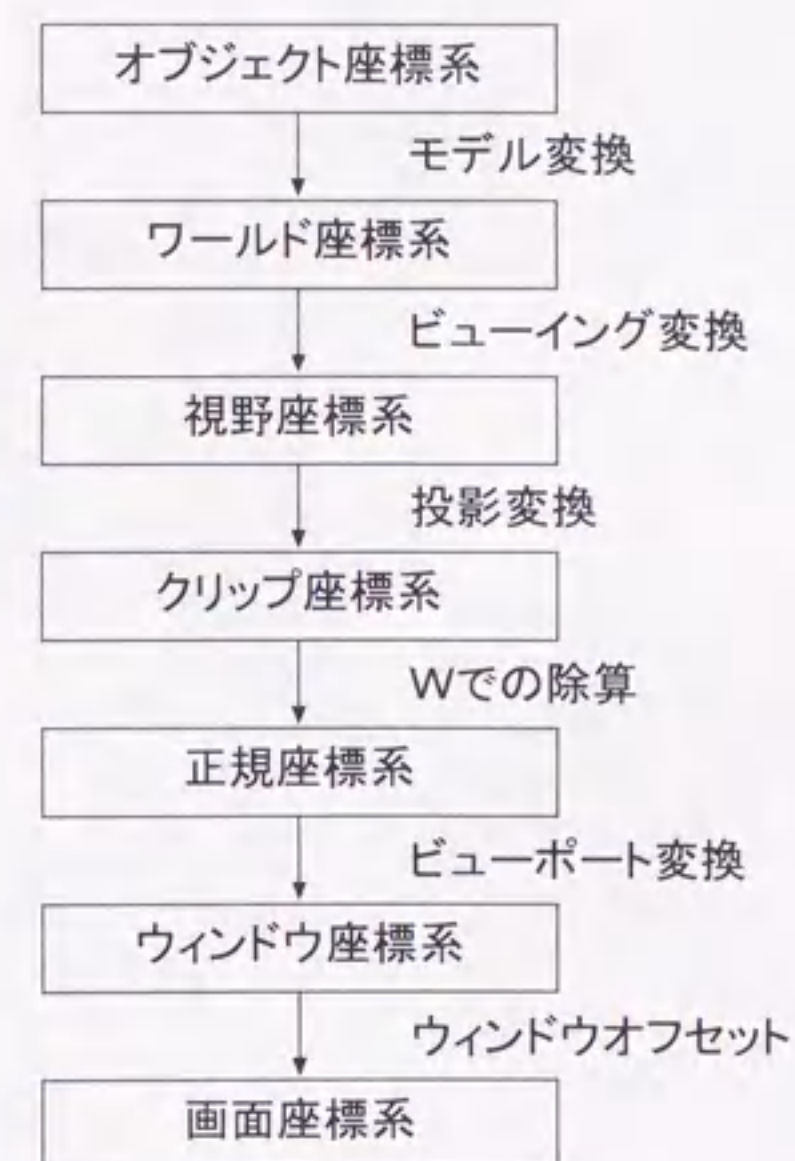


図 3.7: IrisGL 内の座標変換の流れ。



## i) オブジェクト座標系

各モデル固有の座標系である。CGはモデルの集まりで表現されている(図3.8)。

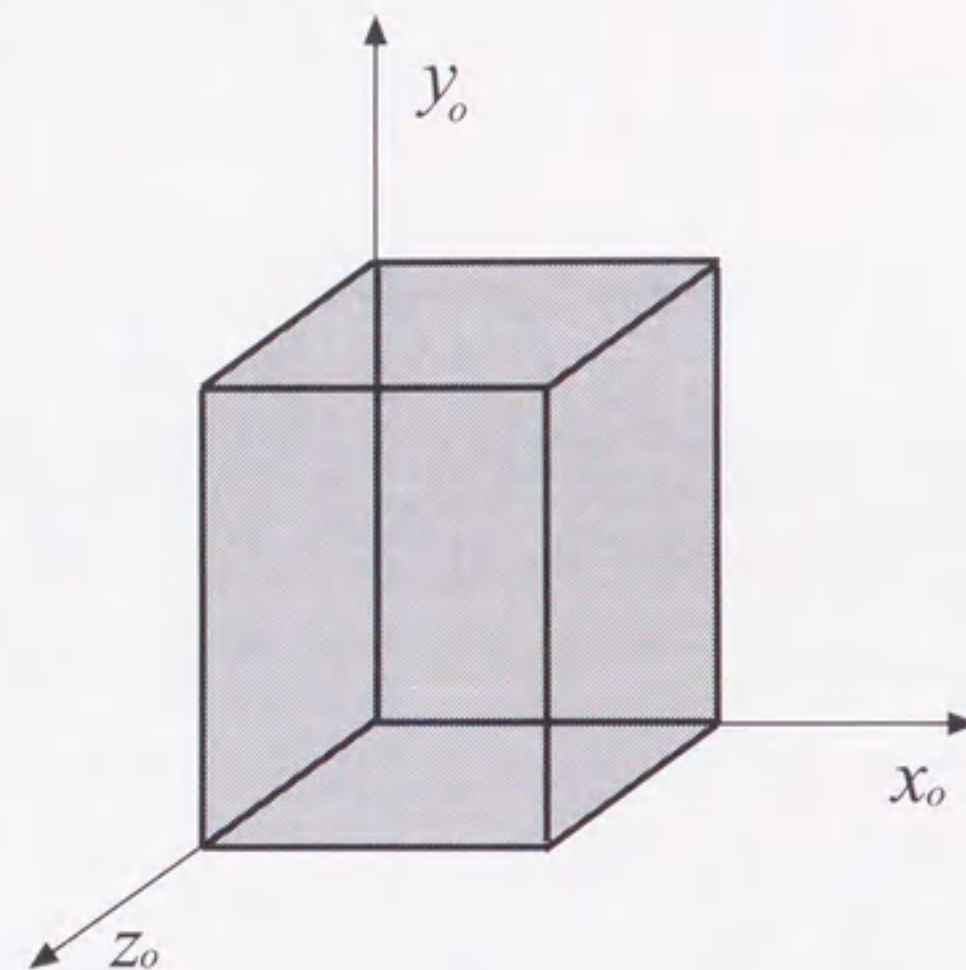


図 3.8: オブジェクト座標系.

## ii) ワールド座標系

各モデルがモデリング変換によって、ワールド座標系の絶対位置に変換される(図3.9). この座標系が仮想空間の座標系である.

## iii) 視野座標系

ワールド座標系がビューイング変換によって、視野座標系に変換される(図3.10). ビューイング変換は視点と視線を指定することによってIrisGL内部で計算される.

IrisGLの投影変換では、視点は原点で、視線は $-z$ 軸方向を向いているものとして処理される. そのため、ビューイング変換では、指定した視点・視線が原点・ $-z$ 軸方向となるように、各モデルの相対関係を保ちながらワールド座標系を平行移動・回転する.

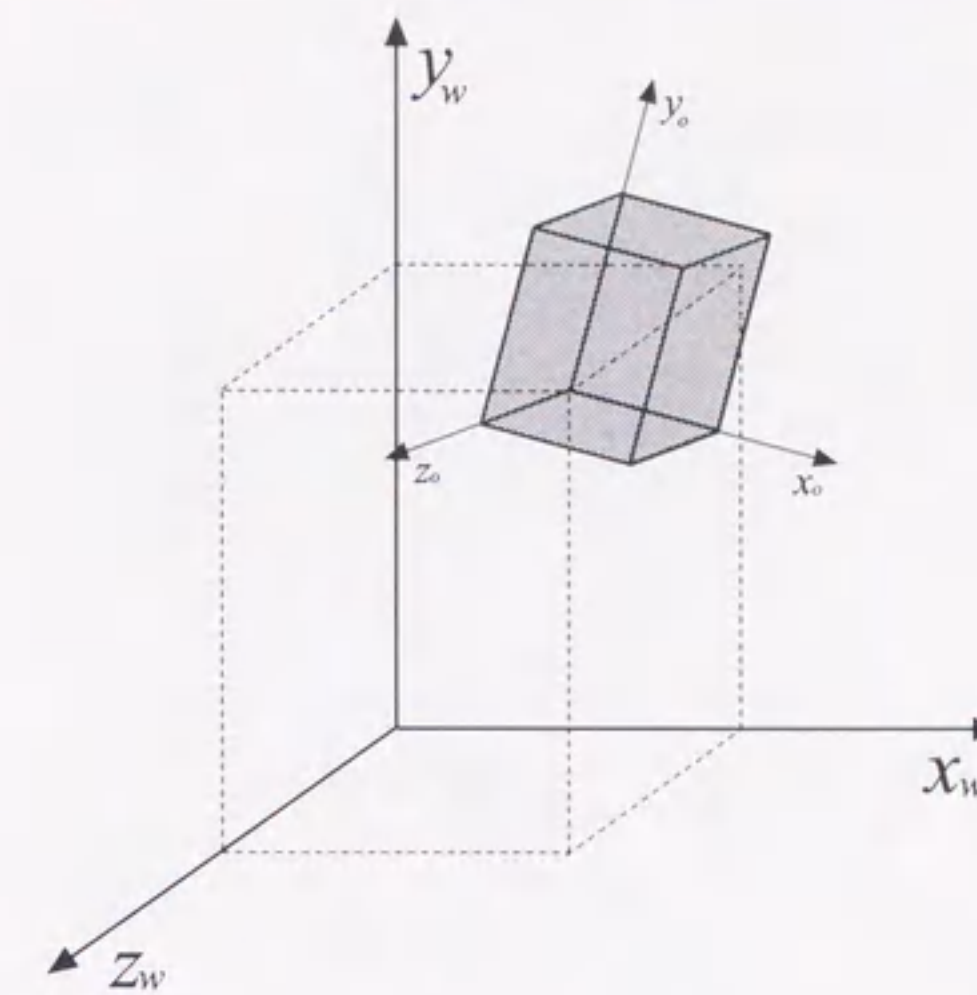


図 3.9: ワールド座標系.

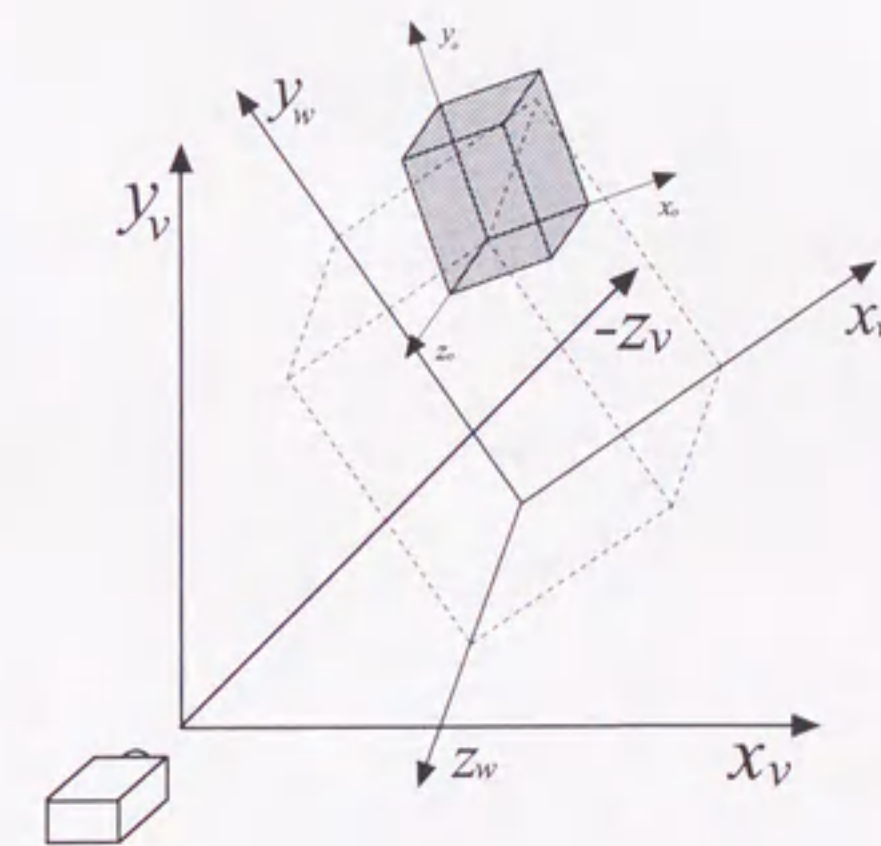


図 3.10: 視野座標系.



## iv) クリップ座標系

視野座標系が後述する視体積によって定義された投影変換によって、クリップ座標系に変換される。投影方法には、透視投影・ウィンドウ投影・正投影を選択する(図 3.11 は透視投影)。

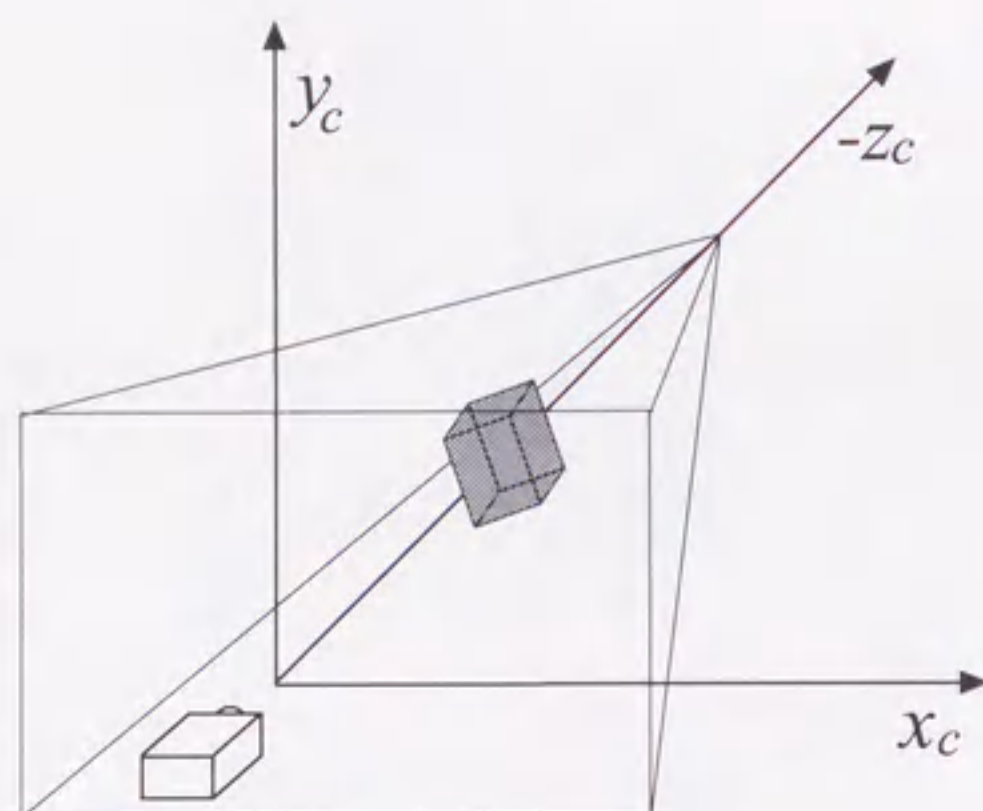


図 3.11: クリップ座標系.

## v) 正規座標系

クリップ座標系に変換されている空間直方体を、各軸の幅  $W$  で除算することによって、単位立方体 ( $-1 \leq x, y, z \leq 1$ ) に正規化する。これは IrisGL の内部処理であるため、ユーザが関与する変換ではない(図 3.12)。

## vi) ウィンドウ座標系

二次元座標に投影される視野座標系の領域を定義するビューポート変換によって、ウィンドウ座標系となる。一般にビューポート変換は 1:1 の変換としている。ウィンドウ座標系は CG を描画するウィンドウに対応する座標系で、ウィンドウの左下端点がウィンドウ座標の  $(0, 0)$  に対応する。正規座標系の  $x, y$  座標を直接二次元にスケーリングした座標系

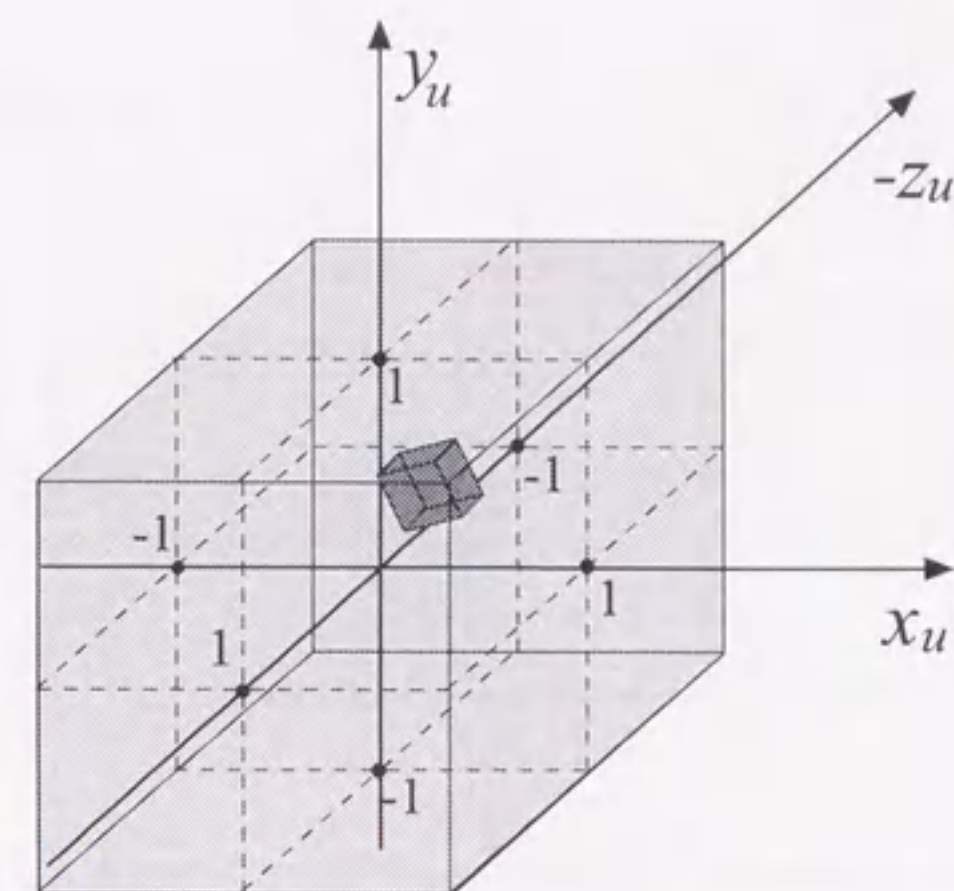


図 3.12: 正規座標系.

である(図 3.13)。実際には三次元的にスケーリングされており、このときの  $z$  値は、隠面消去やデプス・キューイングに使用される。

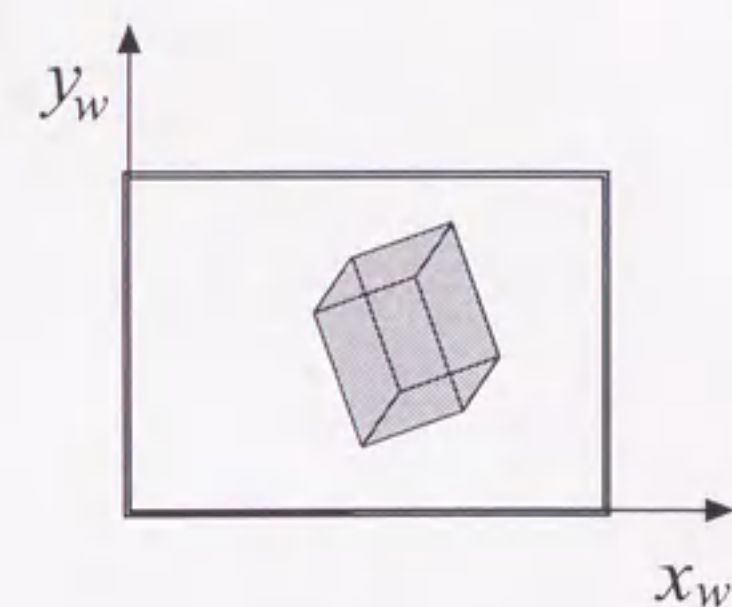


図 3.13: ウィンドウ座標系.

## vii) 画面座標系

画面座標系は実際に描画される画面の座標系で、画面中のウィンドウの位置を示すオフセットによってウィンドウ座標系から変換される。



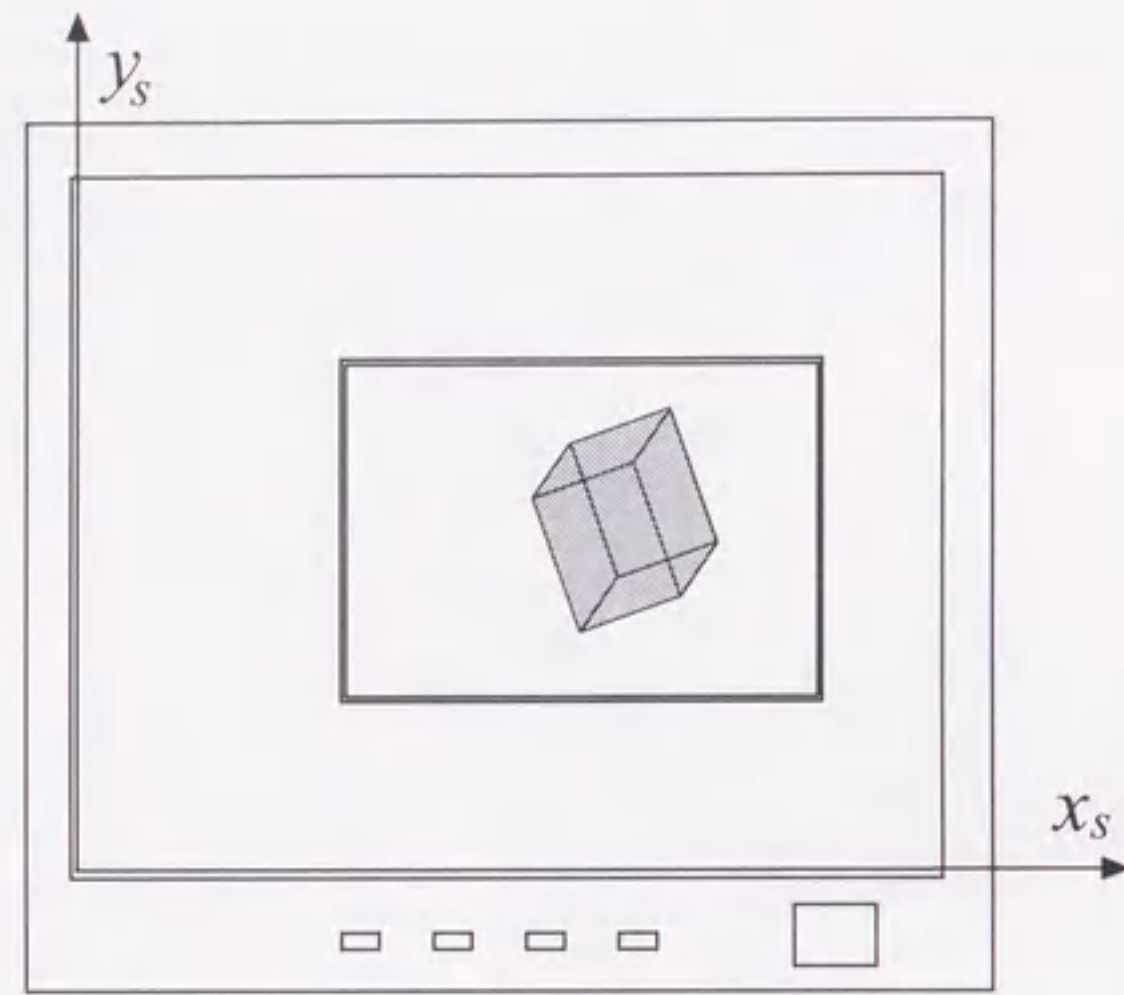


図 3.14: 画面座標系.

以上をまとめると、ワールド座標系から画面座標系へ変換するには、ビューイング変換、投影変換、ウィンドウオフセットを求めなければならない。ビューイング変換には視点位置・視線ベクトル、投影変換には視体積、ウィンドウオフセットには画面の描画位置の情報が必要である。

そこで次節では、キャリブレーションで求められた変換行列から視点位置・視線ベクトル、視体積を導出する方法を述べる。なお、ウィンドウオフセットは、図 3.6 のような SDP-4 のコンフィグレーションで決定する。

### 3.4.3 GL パラメータの算出

#### i) 視点位置の算出方法

視点  $V_m$  の求め方を、図 3.15 を用いて説明する。

まず、画面座標のある点  $P_s(x_s, y_s)$  を与え、ワールド座標の点  $P_w(x_w, y_w, z_w)$  を求める場

### 3.4. 実空間と仮想空間の座標系の整合

合を考える。式 (3.1) において  $h = 1$  とすれば、

$$\begin{cases} T_{11}x_w + T_{21}y_w + T_{31}z_w + T_{41} = x_s \\ T_{12}x_w + T_{22}y_w + T_{32}z_w + T_{42} = y_s \end{cases} \quad (3.3)$$

である。これらの式より求められる直線  $l$  は、透視変換の性質上、視点  $V_w$  を通り画面座標  $P_s$  に投影される直線である。したがって、直線  $l$  は、方向ベクトルを  $d_w$ 、 $k$  を任意の数とすると、

$$P_w = V_w + kd_w \quad (3.4)$$

と表される。

次に、式 (3.1) において  $h = 0$  とすると、

$$\begin{bmatrix} x_w & y_w & z_w & 1 \end{bmatrix} T = 0 \begin{bmatrix} x_s & y_s & 0 & 1 \end{bmatrix} \quad (3.5)$$

である。この時の座標  $P_w$  は、画面座標  $P_s$  に依らず全ての直線  $l$  が通る点、すなわち視点  $V_w(x_{w0}, y_{w0}, z_{w0})$  を示す。この式を展開すると、

$$\begin{cases} T_{11}x_{w0} + T_{21}y_{w0} + T_{31}z_{w0} + T_{41} = 0 \\ T_{12}x_{w0} + T_{22}y_{w0} + T_{32}z_{w0} + T_{42} = 0 \\ T_{14}x_{w0} + T_{24}y_{w0} + T_{34}z_{w0} + 1 = 0 \end{cases} \quad (3.6)$$

であるから、

$$\begin{bmatrix} x_{w0} & y_{w0} & z_{w0} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} & T_{14} \\ T_{21} & T_{22} & T_{24} \\ T_{31} & T_{32} & T_{34} \end{bmatrix} = \begin{bmatrix} -T_{41} & -T_{42} & -1 \end{bmatrix} \quad (3.7)$$

となる。したがって、

$$\begin{bmatrix} x_{w0} & y_{w0} & z_{w0} \end{bmatrix} = \begin{bmatrix} -T_{41} & -T_{42} & -1 \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} & T_{14} \\ T_{21} & T_{22} & T_{24} \\ T_{31} & T_{32} & T_{34} \end{bmatrix}^{-1} \quad (3.8)$$

となる。このようにして視点位置  $V_w$  を求めることができる。



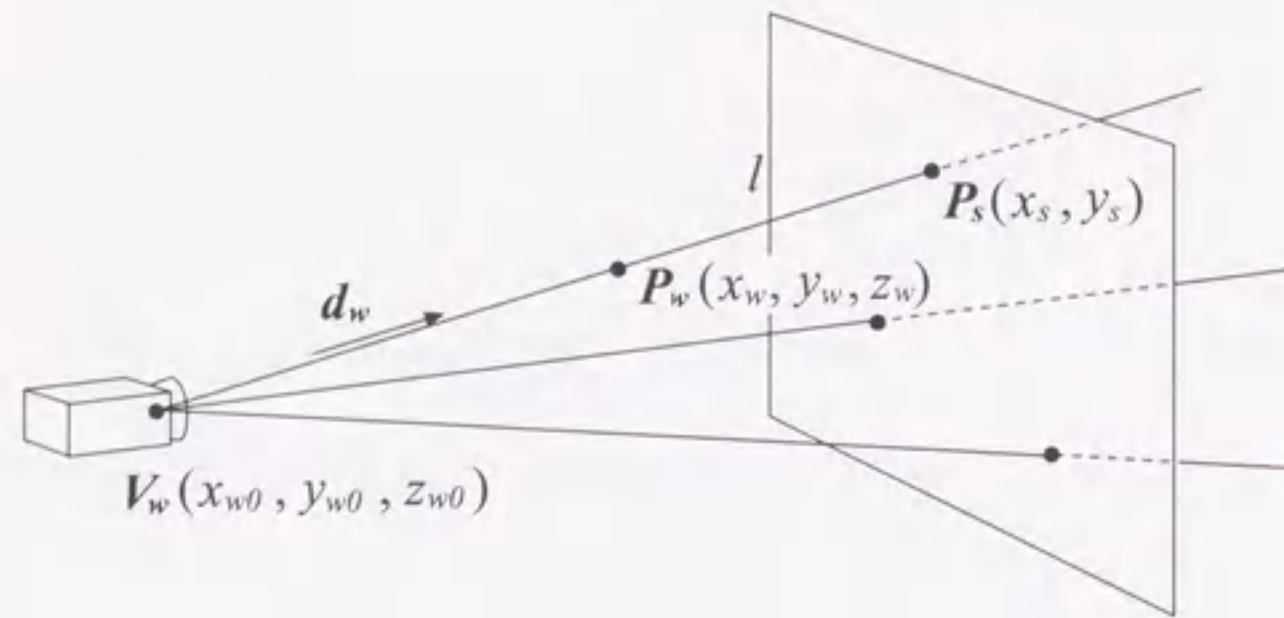


図 3.15: 視点位置の計算.

ii) 視線ベクトルの算出方法

式(3.4)で表される直線  $l$  上の点  $P_w$  を一意に決定するのは、式(3.1)より得られる、

$$T_{14}x_w + T_{24}y_w + T_{34}z_w + 1 = h \tag{3.9}$$

である。図3.16のように、この式は法線ベクトル  $d_h$  が  $(T_{14}, T_{24}, T_{34})$  である平面  $p$  を表し、この平面  $p$  と直線  $l$  の交点によって、 $P_w$  が決定される。この平面  $p$  は視線ベクトル  $v_w$  に垂直であるので、法線ベクトル  $d_h$  の向きは視線ベクトル  $v_w$  の向きと一致する。

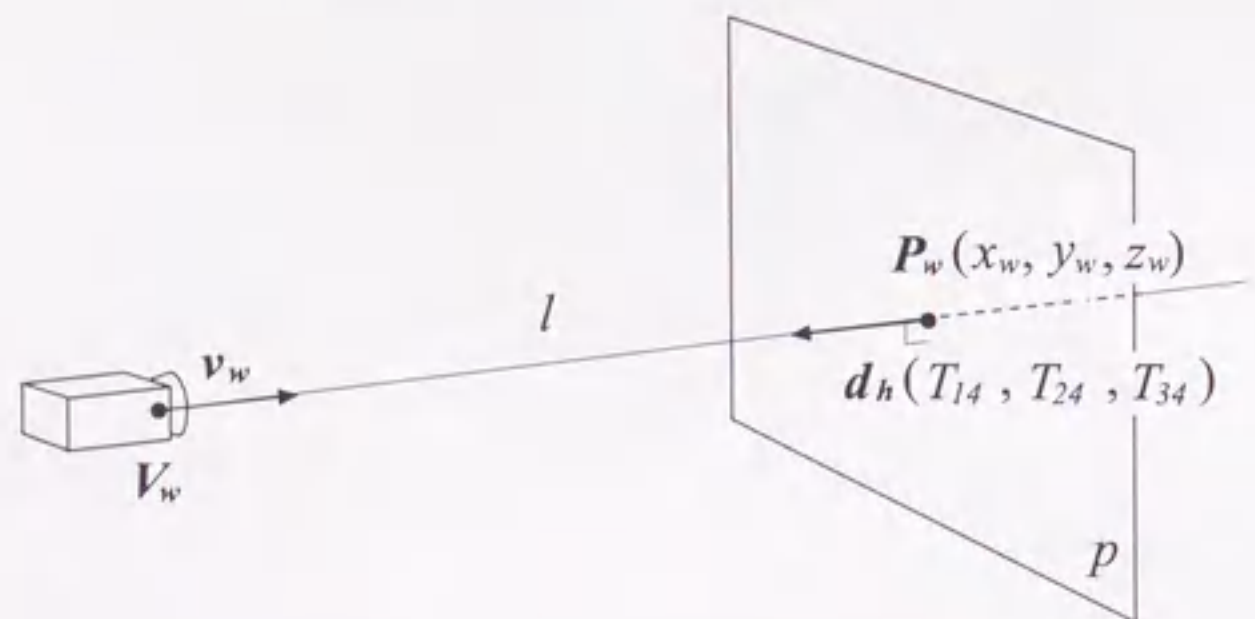


図 3.16: 視点位置の計算.

iii) ウィンドウ投影変換の算出方法

ウィンドウ投影変換は透視投影の一種で、図3.17のように視四角錐によって定義される。一般に透視投影は視線ベクトルに対して対称であるが、ウィンドウ投影では非対称な視角錐体を定義できる。ここでは、合成後にカメラの光軸が映像の中心にない場合を考えて、ウィンドウ投影を用いている。

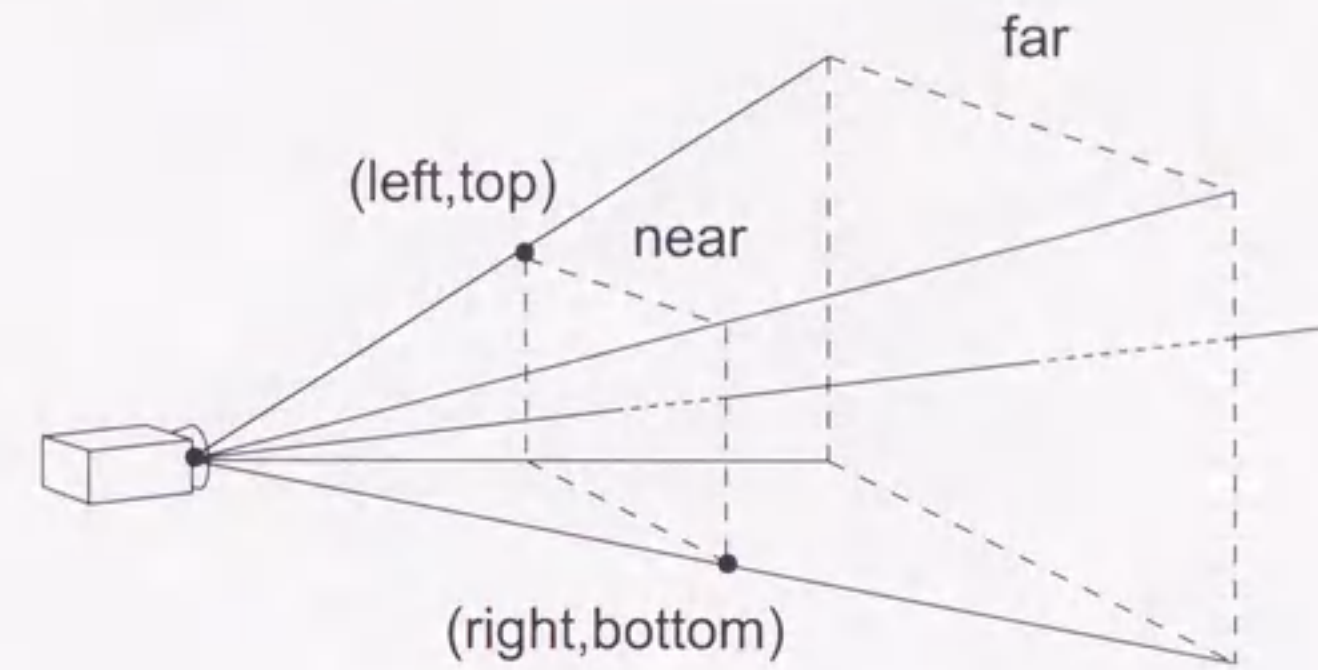


図 3.17: ウィンドウ変換.

同図における (left, top), (right, bottom) は視野座標系で、near 平面での座標である。そして near, far は、仮想空間のスケールによって決められる値で、near 平面から far 平面の間のみ描画される。

図中の視四角錐の各斜辺は、画面座標系の表示範囲矩形の各頂点に投影される直線である。従って、表示範囲矩形の各頂点を  $P_s$  とし式(3.4)を求めると、 $d_w$  はワールド座標系での各斜辺の方向ベクトルとなる。ここで、

$$d'_w = \text{near} \times \frac{d_w}{|d_w|} \tag{3.10}$$

とすれば、 $d'_w$  は大きさ near の方向ベクトルとなる。そして視線ベクトル  $v_w$  を用いて、方向ベクトル  $d'_w$  をワールド座標系から視野座標系  $d'_v$  に変換すると、各斜辺の方向ベクトル  $d'_v$  の  $x, y$  座標は left, top, right, bottom の値となる。



## iv) クリッププレーンの指定

実空間座標系とワールド座標系が一致しているとする、図3.18のように、クリッププレーンはワールド座標系での物体の位置  $P_w$  と向き  $n_w$  で決定される。

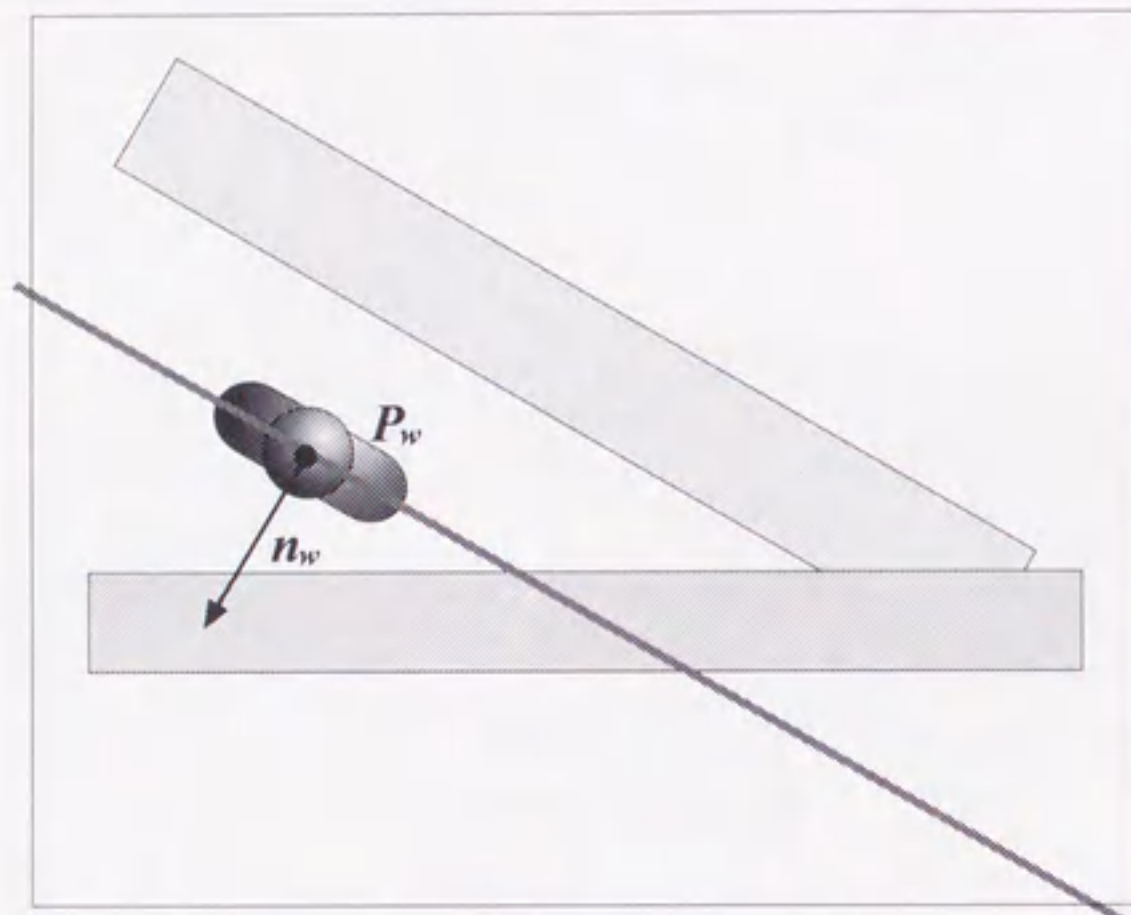


図 3.18: クリッププレーンの計算.

しかし、IrisGLでは視野座標系でクリッププレーンを定義しなくてはならない。そこで、ビューイング変換行列  $M_v$  が与えられているとすれば、視野座標系でのクリッププレーンの位置  $P_v(P_x, P_y, P_z)$  と  $n_v(n_x, n_y, n_z)$  は

$$P_v = P_w M_v \quad (3.11)$$

$$n_v = n_w M_v \quad (3.12)$$

から定まる。したがって、視野座標系でのクリッププレーンの平面方程式は、

$$n_x(x - P_x) + n_y(y - P_y) + n_z(z - P_z) = 0 \quad (3.13)$$

となる。これを整理すると、

$$n_x x + n_y y + n_z z - (n_x P_x + n_y P_y + n_z P_z) = 0 \quad (3.14)$$

である。ここで、平面方程式を、

$$Ax + By + Cz + D = 0 \quad (3.15)$$

とすれば、

$$\begin{bmatrix} A & B & C & D \end{bmatrix} = \begin{bmatrix} n_x & n_y & n_z & - \begin{bmatrix} n_x & n_y & n_z \end{bmatrix} \begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} \end{bmatrix} \quad (3.16)$$

である。このようにして平面方程式の係数を求めて、クリッププレーンを指定する。



### 3.5 実空間と仮想空間の変換のキャリブレーション

Mixed Realityでは、実空間と仮想空間を高度に融合するためには正しく座標変換が必要である。そのために高精度にキャリブレーションする必要があることは、すでに述べた通りである。

本研究のシステムで、キャリブレーションが必要な座標変換には次の3つがある。

- (1) グローバル座標系と実空間座標系の変換
- (2) グローバル座標系と仮想空間座標系の変換
- (3) グローバル座標系センサ座標系の変換

ここで、実空間座標系をグローバル座標系とすれば変換(1)は不要となり、必要なキャリブレーションは、

- (1) 実空間座標系と仮想空間座標系の変換
- (2) 実空間座標系とセンサ座標系の変換

の二つになる。しかし、これらのキャリブレーションを手作業で行うのはオペレータに大変な負担となる。そこで本システムを構築するにあたり、もっとも基本的な変換である変換(1)を画像処理によって自動化している。

#### 3.5.1 キャリブレーション方法

第3.4節で述べたように、キャリブレーションでは実空間の既知の点に対応するスクリーン上の点を示す必要がある。本研究では、既知の点を指定するために、正確に測定された立方体(図3.19)を使用した。この立方体に実空間の既知の座標をマークし、クロマキー合成装置の合成出力映像にあるマーカのスクリーン上の座標を指示する。キャリブレーションには既知の座標は6点以上必要であるので、マーカ指示作業を最低6回繰り返す。この作業を左右両眼について行う。

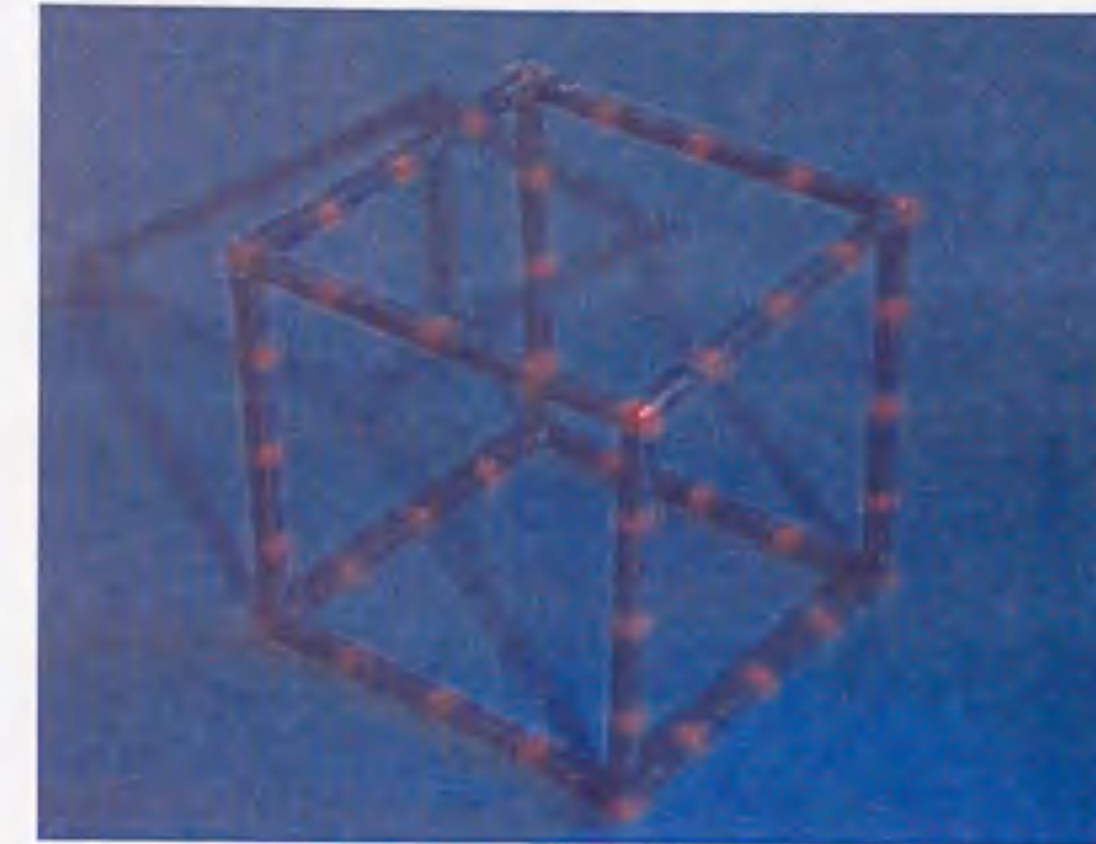


図 3.19: キャリブレーション用立方体 (1 辺 20cm).

#### 3.5.2 手作業によるキャリブレーション

手動キャリブレーションでは、クロマキー合成装置の出力結果をモニターで確認しながらマーカをマウスで指示する(図3.20)。



図 3.20: 手動キャリブレーションの様子。

しかし、クロマキー合成システムには入力出力間に約 300m 秒の遅延があるため、マー



カをマウスで効率的に指示することができない。そして片目の映像のキャリブレーションが終わると、モニタ出力を他片目に切り替える操作を必要とする。そのため非常に作業時間が長く、さらに目に極度な負担がかかる。

### 3.5.3 キャリブレーションの自動化

本研究でのシステムを構築するにあたり、手動キャリブレーション作業のうちマウスでのポインティング作業と、左右両眼映像を切り替える作業を自動化した。このときのシステム構成を図 3.21 に示す。

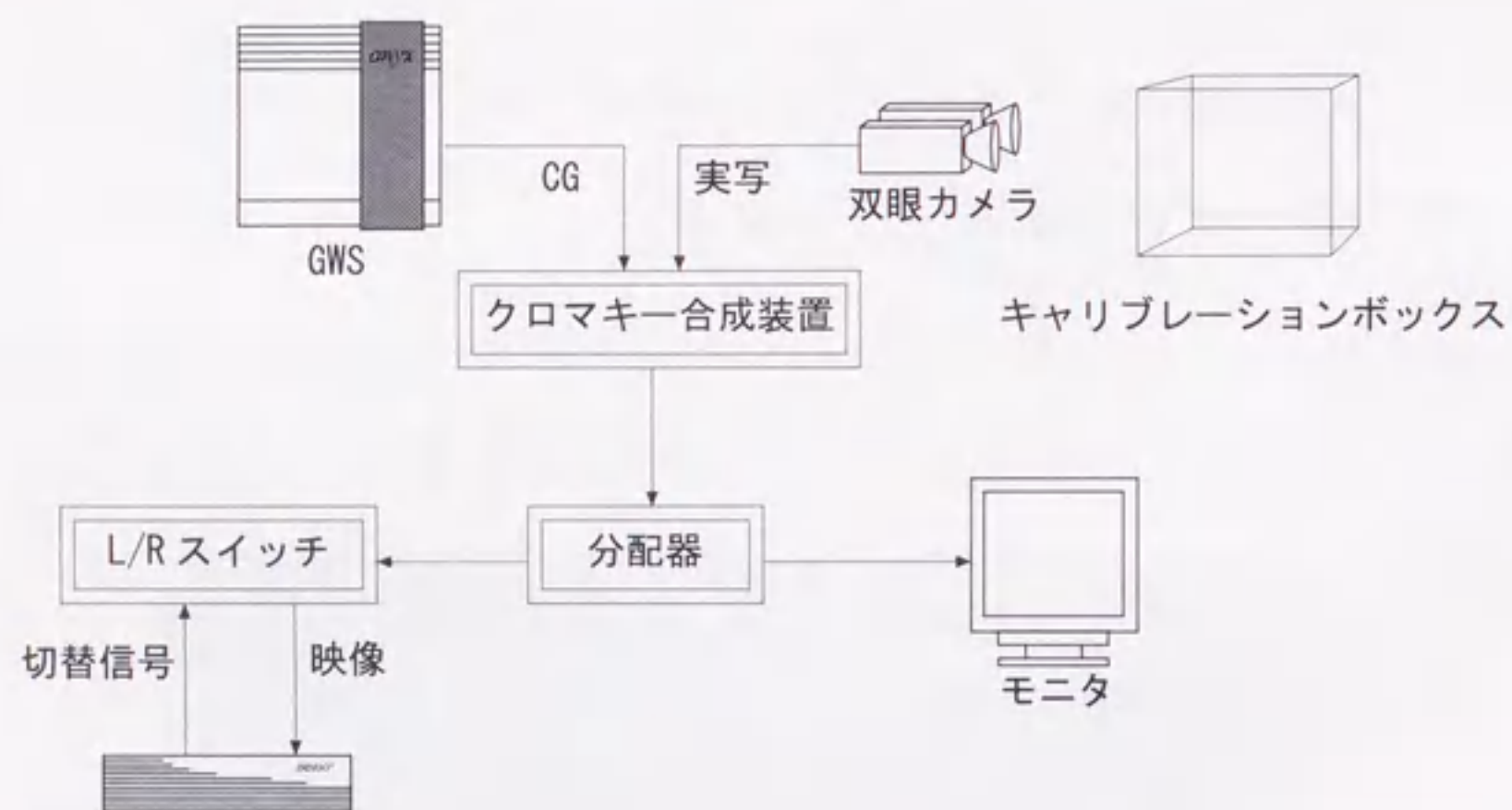


図 3.21: 自動キャリブレーションのシステム構成。

### 3.5.4 自動ポインティング

キャリブレーションボックスのマーカを、画像処理によって自動的に取得するようにした [34][35]。

#### i) 撮影環境と特徴量の決定

対象空間の照明環境は合成時にもっとも適するように設定するため、キャリブレーションボックス撮影時には適していないときがある。例えば、照光量が多すぎると同じ赤色のマーカでも明度がかなり異なる場合には、RGB 標色系では正しく検出できない。

実際にはこのような場合が多いため、明度の変化の影響を受けることなくマーカを検出するアルゴリズムが必要である。そのため、入力画像は RGB 表色系であるが、マーカ検出には HSV 表色系の色相を用いている。

#### ii) マーカの検出

図 3.22 マーカ検出の処理の流れを示す。

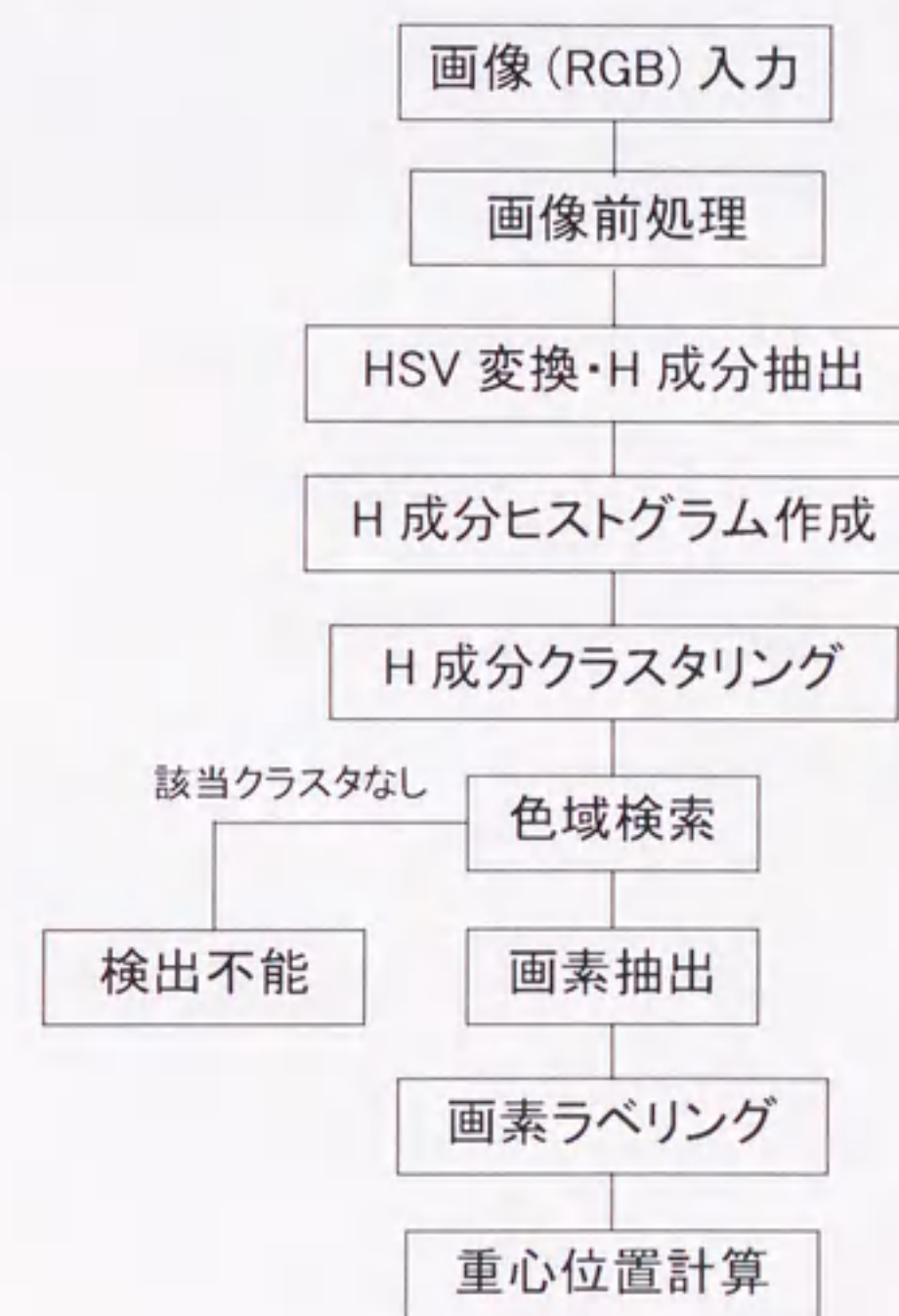


図 3.22: 画像処理の流れ。



## (1) 画像 (RGB) 入力

クロマキー合成装置の出力を Indigo2 HighImpact with IMPACT Compression (表 3.6) に入力して得られた 646×486, RGB 各 8bit の画像である (図 3.23).

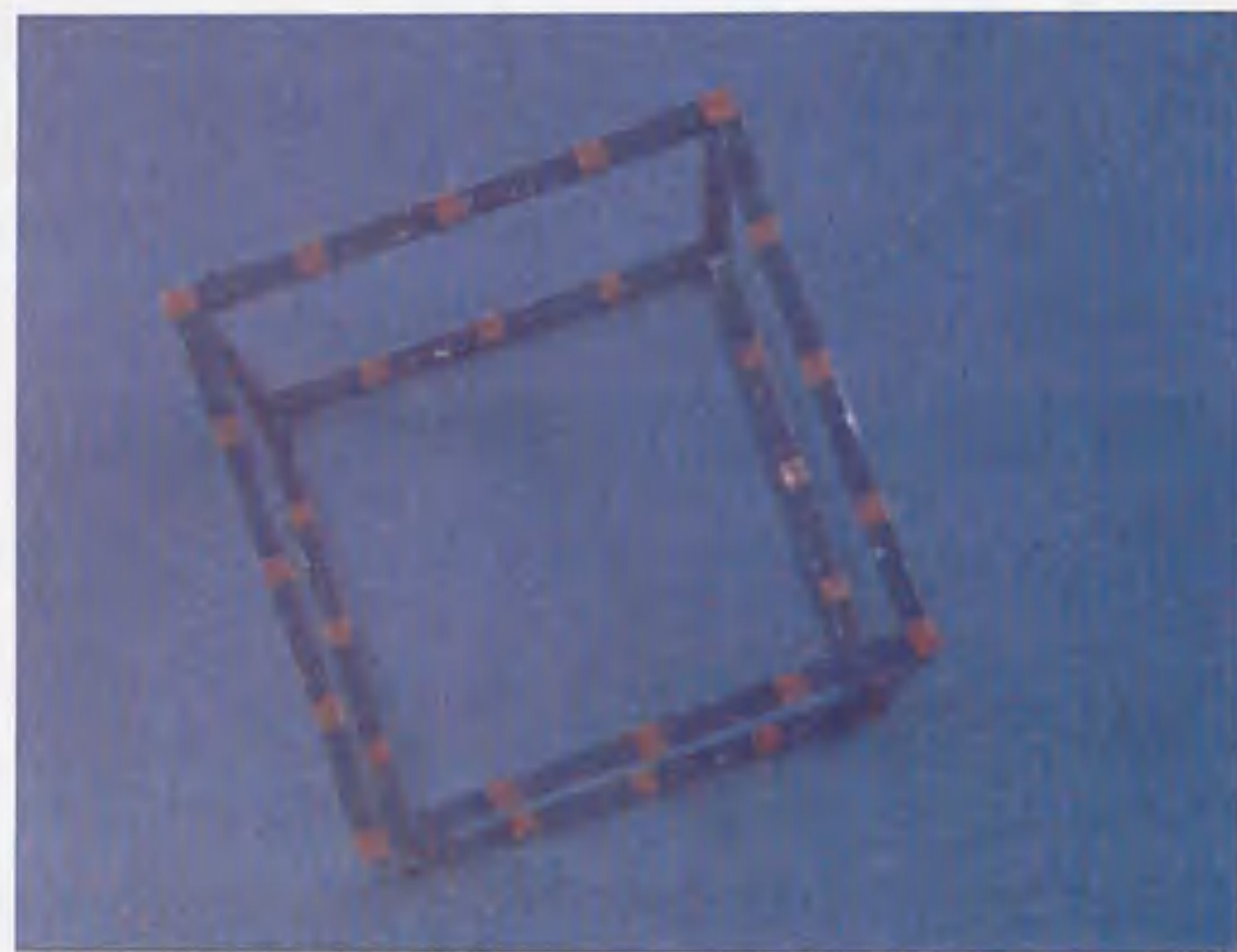


図 3.23: RGB 入力画像.

表 3.6: Indigo2 HighImpact の仕様

項目	仕様
メーカー	Silicon Graphics, Inc.
CPU	MIPS R10000 195MHz
FPU	MIPS R10010
メモリ	128MB
グラフィックボード	High Impact / TRAM
ビデオ入出力ボード	IMPACT Compression (コンポジット, S-Video 各1入力1出力)
OS	IRIX 6.2

## (2) 画像前処理

平滑フィルタによるノイズ除去処理を行う.

## (3) HSV 変換・H 成分抽出

RGB 画像を式 (3.17) で表される Raines の定義による HSV 変換をする [36]. Raines の定義では, RGB は 0 から 1 で, HSV 各成分はそれぞれ  $0 \leq H < 2\pi$ ,  $0 \leq S \leq 2/3$ ,  $0 \leq V \leq 3$  に規格化されている. ただし, 明度が大きいとき (ハレーションなど) では正しく色相が求まらないことがあるので,  $S < 0.02$  では NO HUE とした. 図 3.24 は,  $S = 2/3$ ,  $V = 3$  とした色相画像である.

$$\begin{cases} H = \tan^{-1} \frac{G-B}{2R-G-B} \\ S = \frac{(B-R)^2 + (R-G)^2 + (G-B)^2}{3} \\ V = R + G + B \end{cases} \quad (3.17)$$

なお, HSV→RGB 変換は,

$$\begin{cases} R = \frac{V}{3} + 2S \cos \frac{H}{6} \\ G = \frac{V}{3} - S \cos \frac{H}{6} + S \sin \frac{H}{2} \\ B = \frac{V}{3} - S \cos \frac{H}{6} - S \sin \frac{H}{2} \end{cases} \quad (3.18)$$

である.

## (4) H 成分ヒストグラム作成・クラスタリング

抽出する色相の範囲を選択するために, まず H 成分のヒストグラムを作る. そして, ヒストグラムのピークを利用してクラスタリングを行う. なお, ピークの高周波成分を除去するために, ヒストグラム画像に対しての膨張処理をしている. 図 3.25 は H 成分のヒストグラムで, クラスタはグラフ下方の白線で表している.

## (5) 色域検索

目的の色相値がどのクラスタに所属するかを検索する. どこにも所属していないときは検出不能とする.

## (6) 画素抽出・ラベリング・重心位置計算.

入力画像の中から色相クラスタに所属している画素を抽出し, 抽出された画素に対し



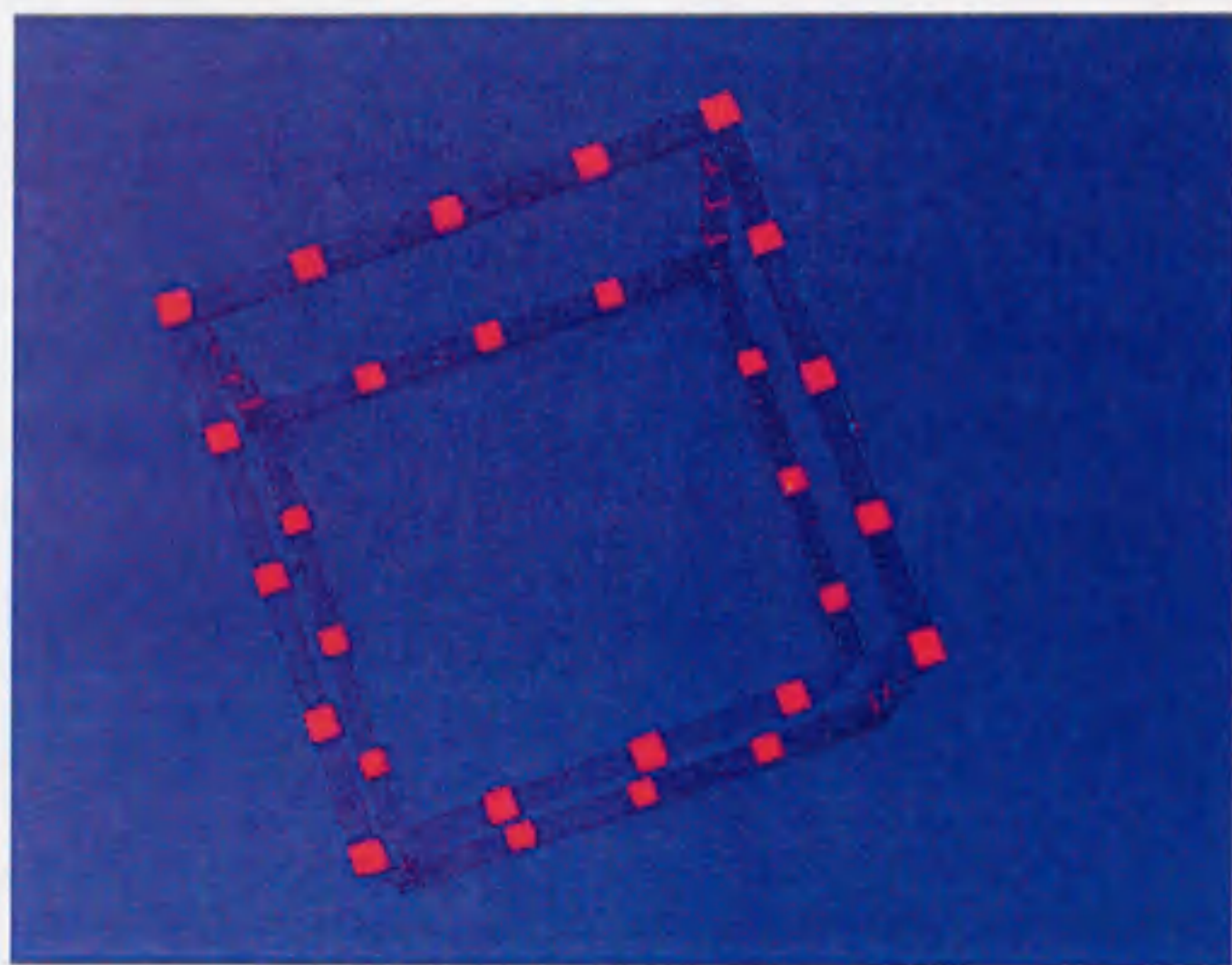


図 3.24: 色相画像.

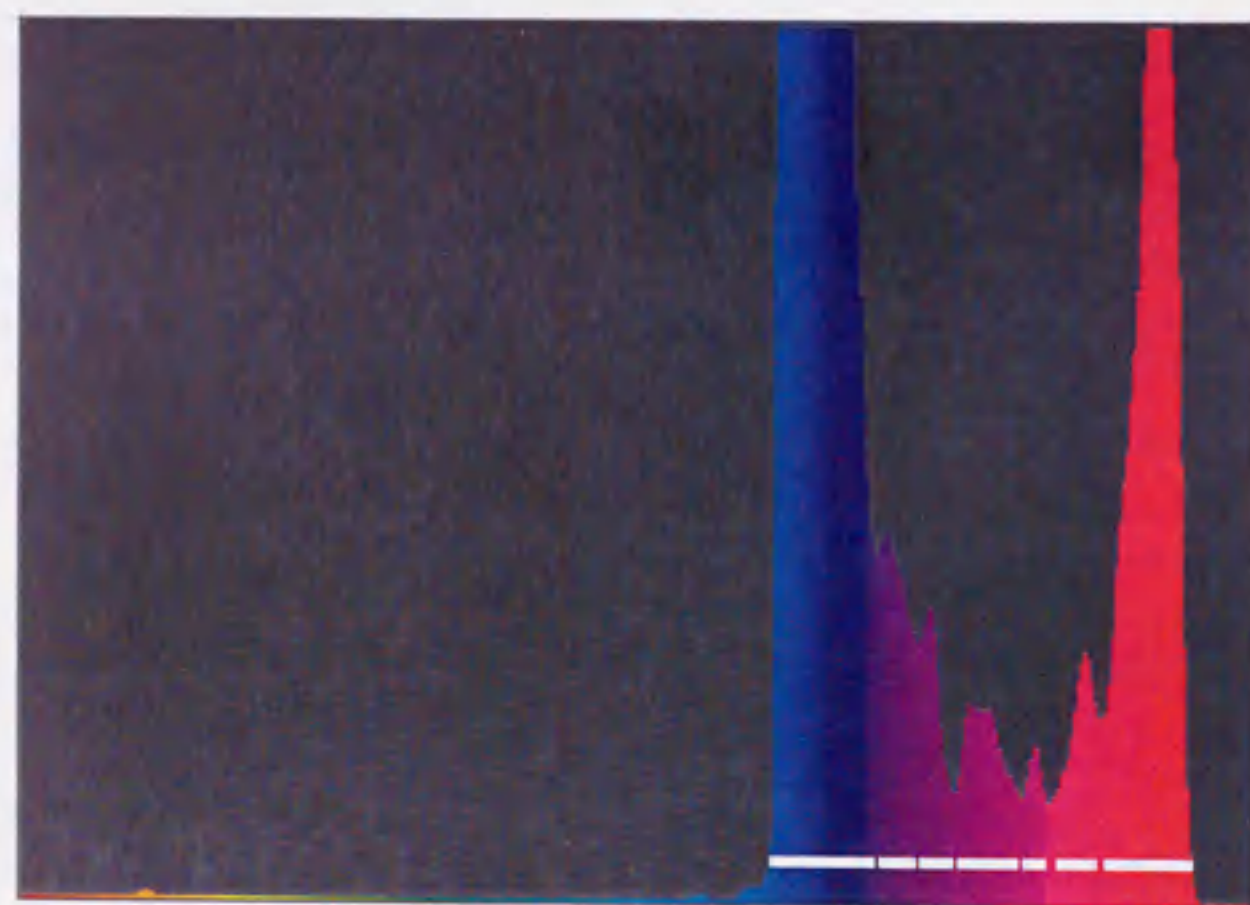


図 3.25: ヒストグラム.

てラベル付けをする(図3.26),そして各ラベルに対して重心位置を計算し,目的のマーカの検出結果とする.



図 3.26: ラベリング.

### 3.5.5 多点でのキャリブレーション

キャリブレーションには最小二乗法を用いているので,測定点の数が増えれば精度の向上が期待できる.しかし,画像処理で測定点を求める場合,特徴量(色相値など)の分解能をあげる必要があるが,撮影環境や入力画像の品質の点から分解能を上げるのは困難である.例えば,256ポイント(8bit)の特徴量で,20種類(20色)検出しようとする,1種類(1色)あたり15ポイント分しかない.従って少ない種類(色)を扱うことが望まれる.

このような理由から,少ないマーカ色数で多くの点数をキャリブレーションに利用する方法を述べる.

- (1) キャリブレーションボックスの映像を取り込む.



- (2) 測定点のマークを検出する。
- (3) 既知の6点を検出し、カメラパラメータを計算する。
- (4) カメラパラメータから7点目以降の測定点を推測する。
- (5) 推測が正しければカメラパラメータを再計算する。

この方法の特徴は、7点目以降を6点目までで求めたカメラパラメータから推測して自動的にカメラパラメータを再計算することである。このため、測定点が何点あっても必要な色数は7色である。

### 3.5.6 精度の検証

図3.27は多点によるキャリブレーションによる誤差と標準偏差のグラフである。この実験では、1辺60cmのキャリブレーションボックスをピクセルサイズ646×486撮影してキャリブレーションを行っている。このとき、キャリブレーションボックス中心付近のピクセル解像度は約1.7mmである。全体的な傾向として測定点数が増えると誤差の平均と標準偏差が良くなっていくのが分かる。また、23点目で誤差が増えているのは、キャリブレーションボックスの23点目が大きくずれているためと思われる。

### 3.5.7 左右両眼映像の自動切り替え

左右映像の切り替え作業は図3.28のような装置で自動化している。Indigo2やOnyxに装備されているセントロニクス仕様のパラレルポートを使用して左右切り替え信号を出力し、ビデオスイッチで左右を切り替える。

これにより、キャリブレーション映像をコンピュータから選択できるようになった。

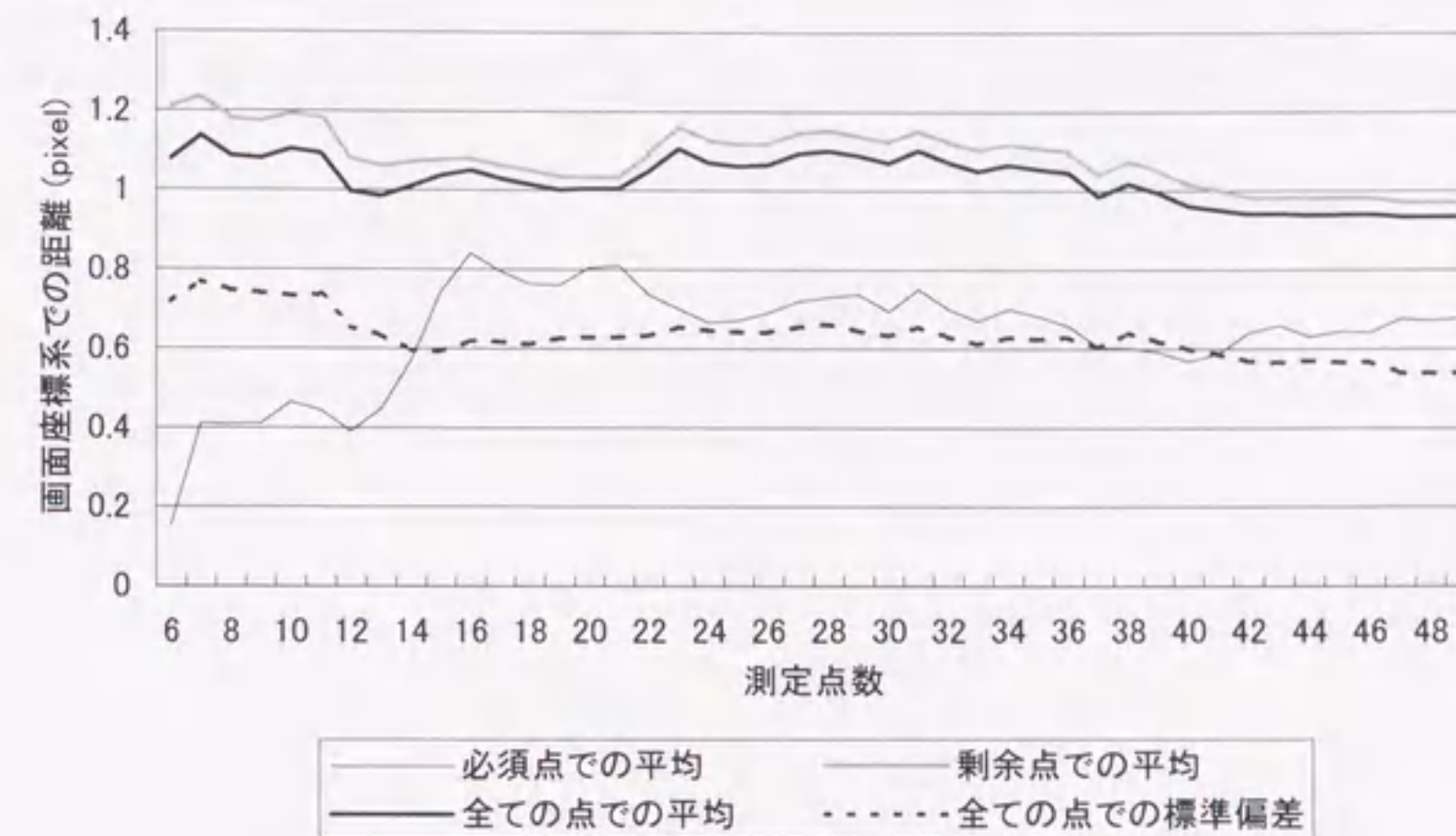


図 3.27: 自動多点キャリブレーションの誤差と標準偏差。

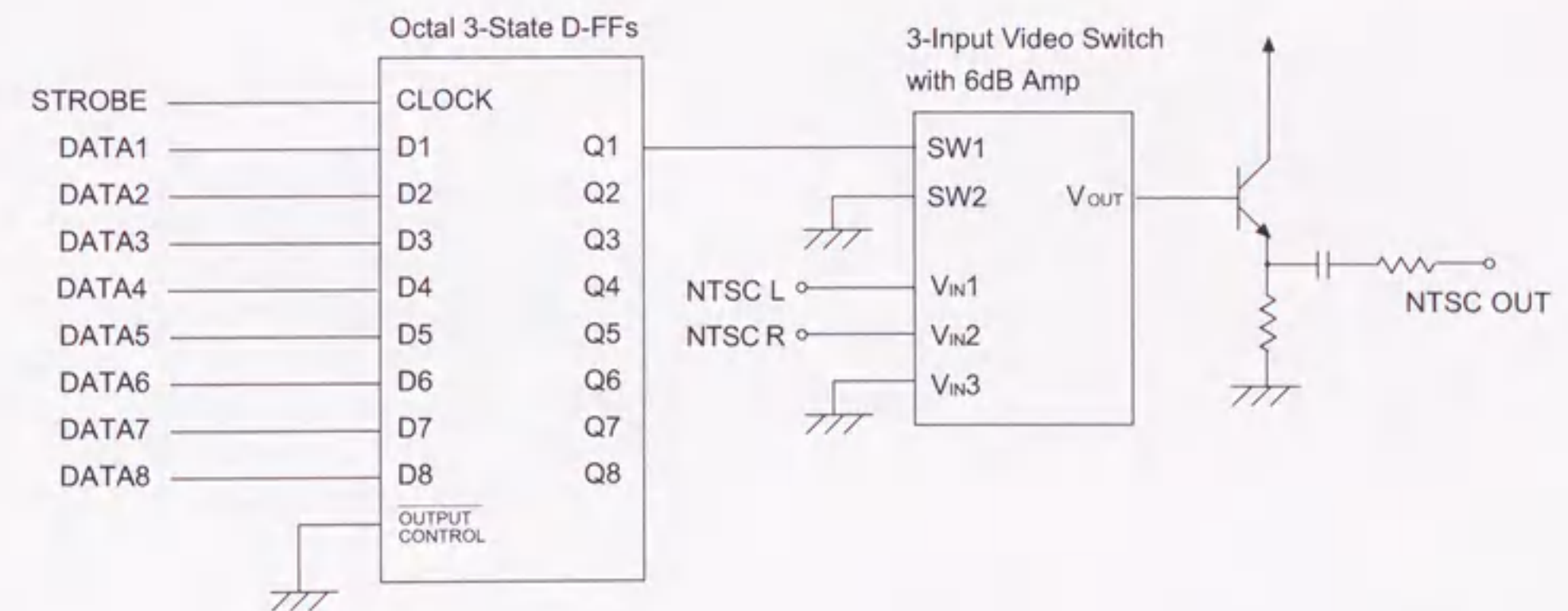


図 3.28: 左右映像切り替え装置。



## 第4章 実空間と仮想空間の合成例

本章では、これまでに述べた方法を用いた合成例を挙げる。本研究では、

- 必要な層の削減

実空間の物体の体積を無視できるならば、斜面3層分割でも2体以上の実物体を仮想空間に挿入できる。このような環境として、スタジオのように広い空間の中の人物を想定して、必要な層の削減の実験を行う。

- 連続的な物体との合成

二つの仮想空間にまたがる連続した仮想物体に、平面的な実物体を挿入する実験を行う。この実験は、実物体の形状を無視できない環境として、卓上での手作業のように、手やエンドエフェクタの形状が重要な場面を想定する。

の2つの環境で実験を行った。

### 4.1 必要な層の削減例: バーチャルスタジオ

バーチャルスタジオとは、舞台セットをCGで作成し登場人物と合成することによって、舞台セットの制作コストを省いたり、実際には人が居られない場面に人を登場させたりすることができるスタジオである [37][38][39]。

#### 4.1.1 システム構成

図4.1のように、スタジオの空間上部のCCDカメラによって登場人物の位置を検出し、双眼カメラで撮影した人物をCG空間に合成し、プロジェクタで出力することによって、仮



仮想空間の中に人が入り込むような状態を作り出す実験を行った。

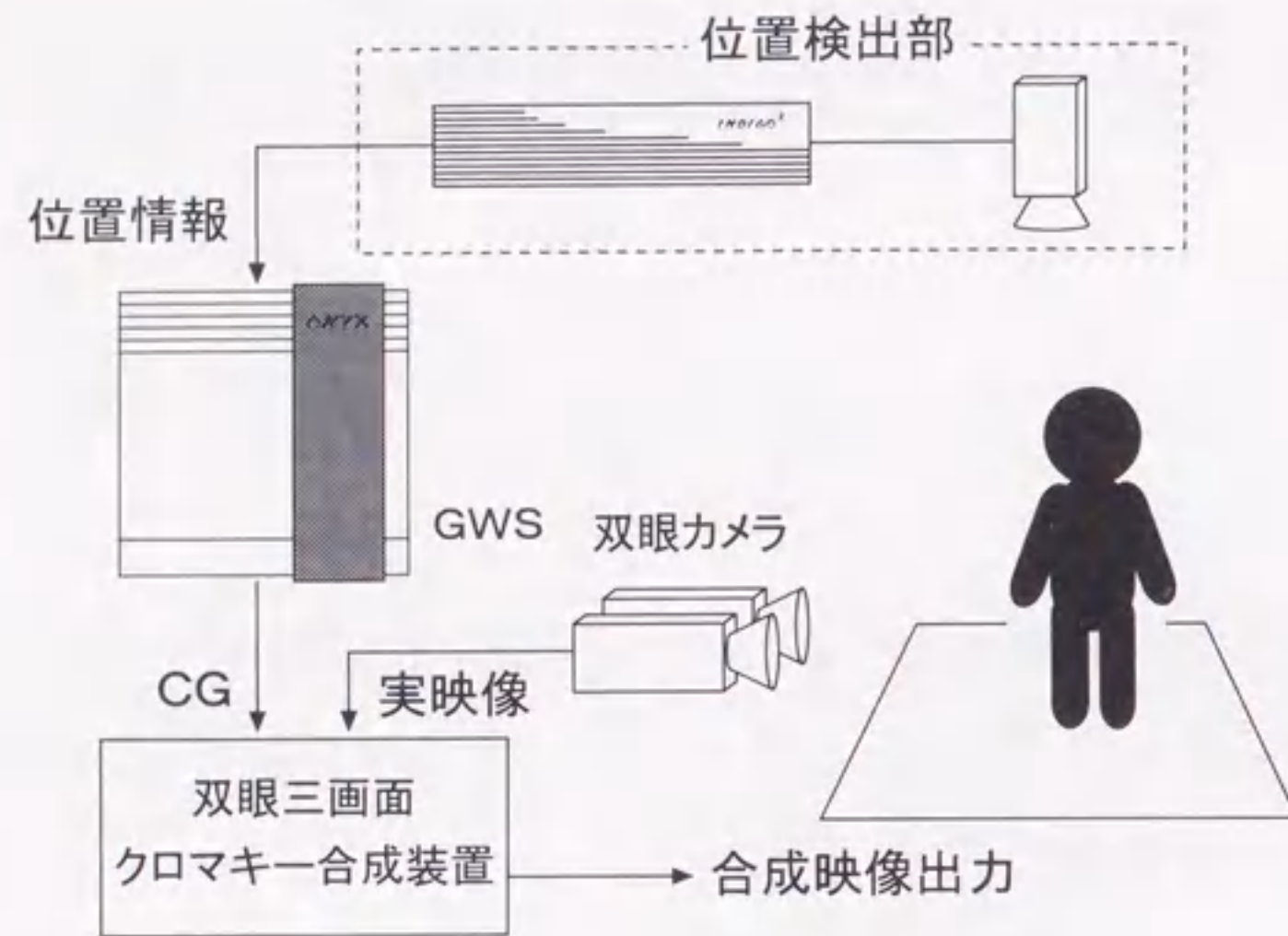


図 4.1: スタジオを使った実験.

#### 4.1.2 CCD カメラによる位置検出

バーチャルスタジオの実験では、カメラ映像の画像処理によって対象物の位置を算出している [40][41].

ここで求められた位置は、クリッププレーンの決定に使われるため、図 4.1 のように対象空間の上部に一台のカメラを設置し、 $(x, y)$  を求めるだけでよい。さらにカメラを使う方法は、実空間の物体にセンサを取り付ける必要がないという利点がある。

##### i) 位置検出アルゴリズム

位置検出アルゴリズムには、合成対象となる実空間が空の映像 (初期画像) と検出対象物体が存在する映像 (検出対象画像) との差分をとり、差分画像の重心位置を対象物体の位置とする方法を用いている。このアルゴリズムは簡単であるため、処理が高速であるという

#### 4.1. 必要な層の削減例: バーチャルスタジオ

利点がある。ここで処理の流れを図 4.2 に示し説明する。

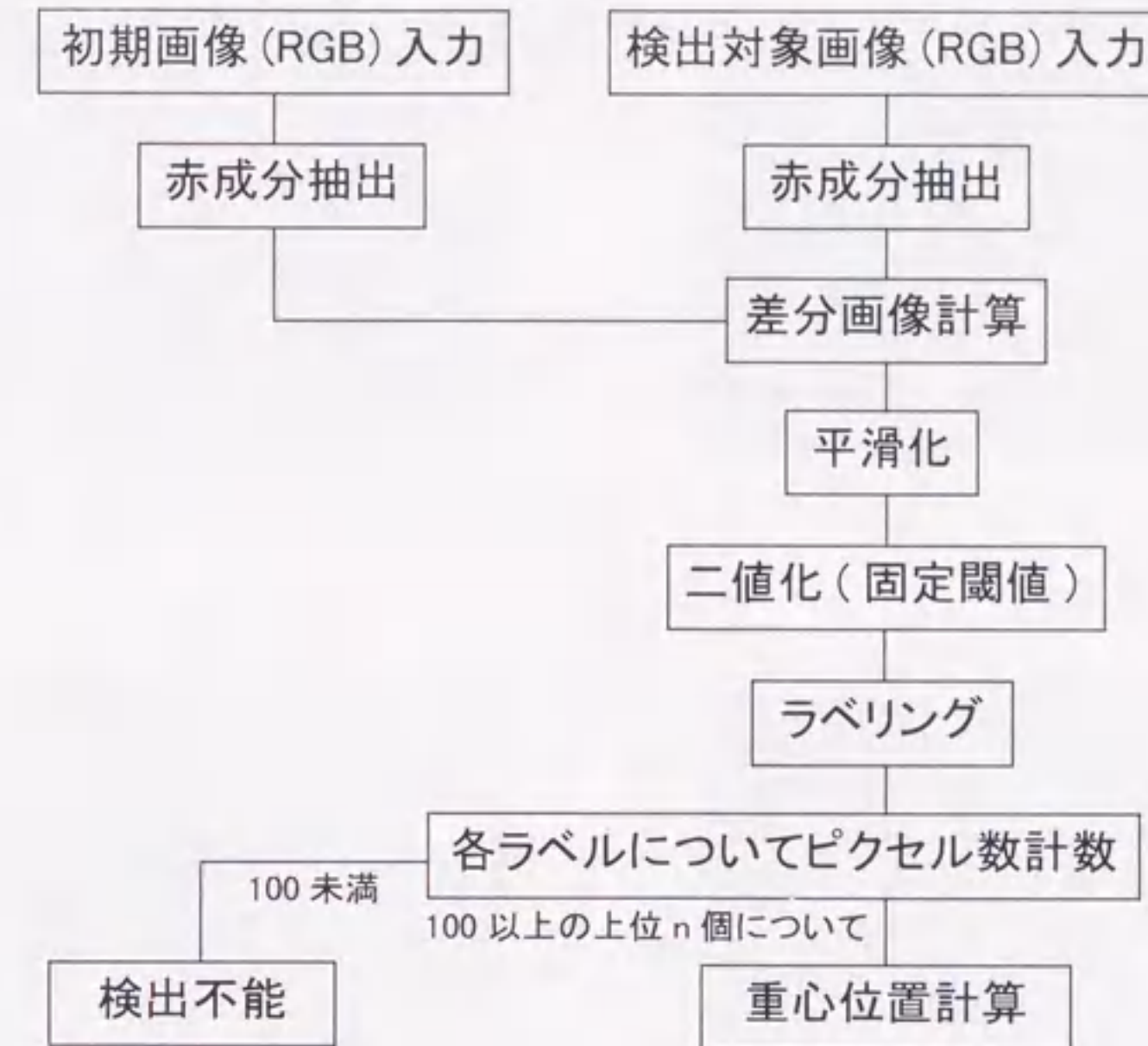


図 4.2: 検出処理の流れ.

##### (1) 初期画像入力・検出対象画像入力・赤成分抽出

対象空間の上方約  $4m$  に広角カメラを設置し、そのカメラで撮影された約  $5m \times 4m$  の矩形領域の映像を Indego2 High Impact with IMPACT Compression でサンプリング ( $646 \times 486$ , RGB 各 8bit) する。

初期画像は検出処理の最初に一度だけ取得し、検出対象画像は検出処理の度に取得する (図 4.3 (a)). そして、処理の高速化のため入力した RGB 画像の赤成分のみを抽出し処理画像とする (図 4.3 (b)). 赤成分を使用したのは、背景 (床) には赤成分が少ないため、赤成分の多い人の肌を検出しやすいと考えられるからである。

##### (2) 差分画像計算

初期画像 (図 4.3 (c)) と検出対象画像の赤成分画像 (図 4.3 (d)) の各ピクセルごと



に差をとり、その差の絶対値を特徴量とした差分画像（図4.3（e））を生成する。

(3) 平滑化・二値化

ノイズ除去のために差分画像を平滑化し、特徴量 256 段階中 176 を閾値としてを二値化をする（図4.3（f））。

(4) ラベリング・重心位置計算

二値化画像に対してラベリング処理をすることによって、検出対象物の候補領域を抽出する。ラベルのピクセル数が 100 ピクセル未満であれば、ノイズ成分と見なし破棄している。100 ピクセル以上のラベルを検出対象物とし、そのラベルの重心位置を計算して検出結果とする。

ii) キャリブレーション

前節の処理で求められた取り込み映像上の重心位置を、実空間での位置に変換するためには、その座標変換式をあらかじめ求めておく必要がある。

今回の実験では、カメラの光軸が対象空間の中心で鉛直であり光学歪みは無視できると仮定することによって、取り込み映像の座標系と実空間の座標系の変換は線形であると見なす。従って、実空間の既知の2点の測定によるキャリブレーションとすることにより簡略化した。

図4.4において、床上の既知の2点  $A(x_{g1}, y_{g1})$ ,  $B(x_{g2}, y_{g2})$  が<sup>s</sup>、取り込み映像の座標  $P(x_{s1}, y_{s1})$ ,  $(x_{s2}, y_{s2})$  にそれぞれ投影されているとする。このとき、検出対象物  $P(x_{gp}, y_{gp})$  が取り込み映像の座標で  $(x_{sp}, y_{sp})$  であったと分かったとすると、点Pの床上の座標は次の式で求められる。

$$x_{gp} = \frac{(x_{sp} - x_{s2})x_{g1} + (x_{s1} - x_{sp})x_{g2}}{x_{s1} - x_{s2}} \quad (4.1)$$

$$y_{gp} = \frac{(y_{sp} - y_{s2})y_{g1} + (y_{s1} - y_{sp})y_{g2}}{y_{s1} - y_{s2}} \quad (4.2)$$



(a) 初期画像



(b) 初期画像赤成分



(c) 検出対象画像



(d) 検出対象赤成分



(e) 赤成分差分画像



(f) 赤成分二値化画像

図 4.3: 検出処理中の画像.



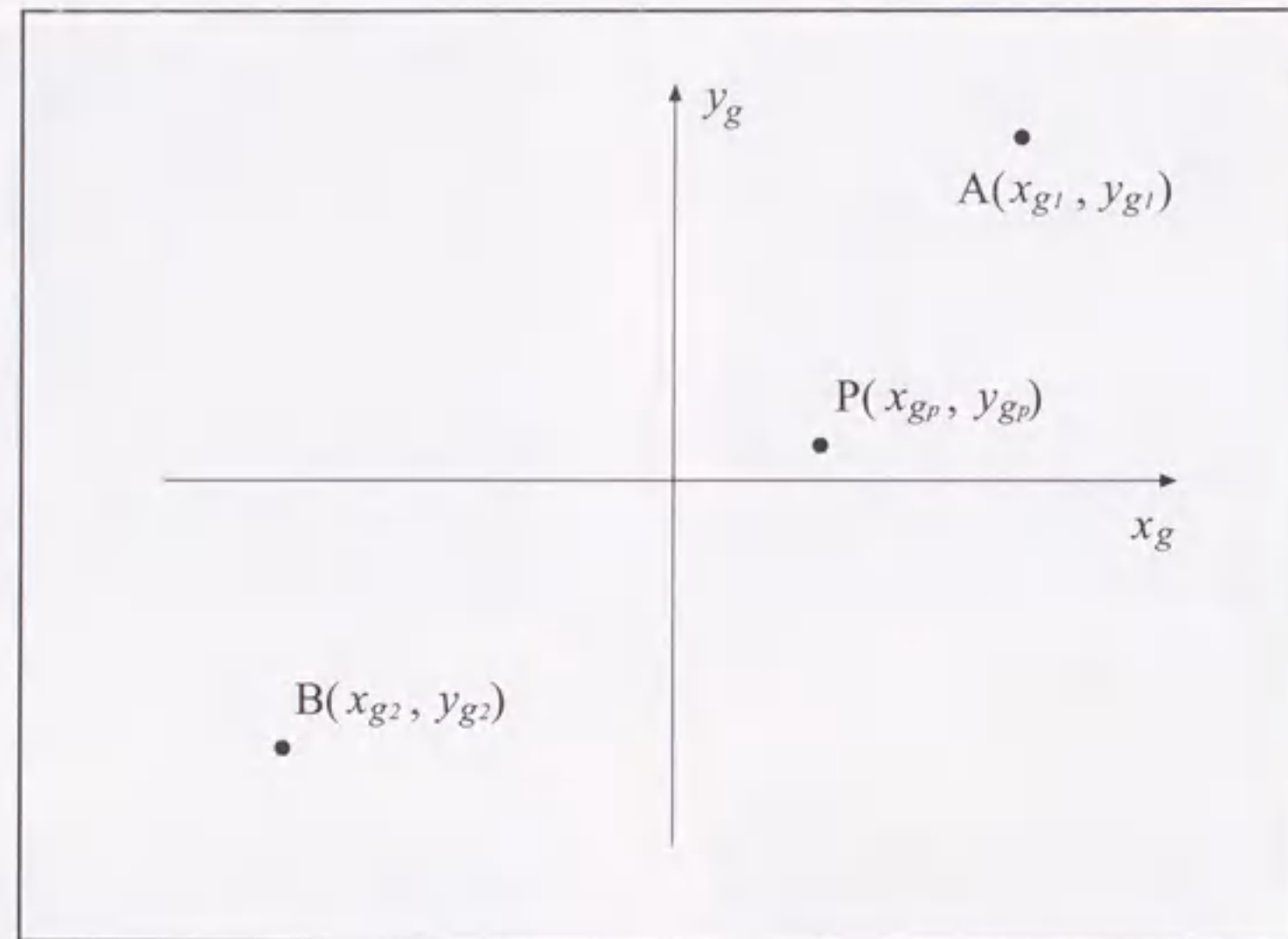


図 4.4: キャリブレーションの方法.

## iii) 位置検出の評価

位置検出精度を評価するため、床に  $3m$  四方の正方形を想定し、キャリブレーションは 1 組の向かい合う頂点で行い、その対角線上を移動した時の位置を測定した。

図 4.5 は測定結果で、 $20cm$  四方の板を床において測定した場合と、クロマキー合成をするときの照明条件で実際に人が往復して歩いたときの場合である。板の場合はほぼ正確に測定できている。人物で測定したときは、軌跡はほぼ対角線上にあるが、波をうった形をしている。これは照明で生じる人物の影の影響と考えられる。

この位置検出方法では、簡略化の仮定から高精度を要求することはできない。しかし簡略化した結果、毎秒約 28 回の測定速度を得ることができた。

これらの結果からスタジオでのクリッププレーンの移動という目的を達成することができる検出方法であると考えられる。

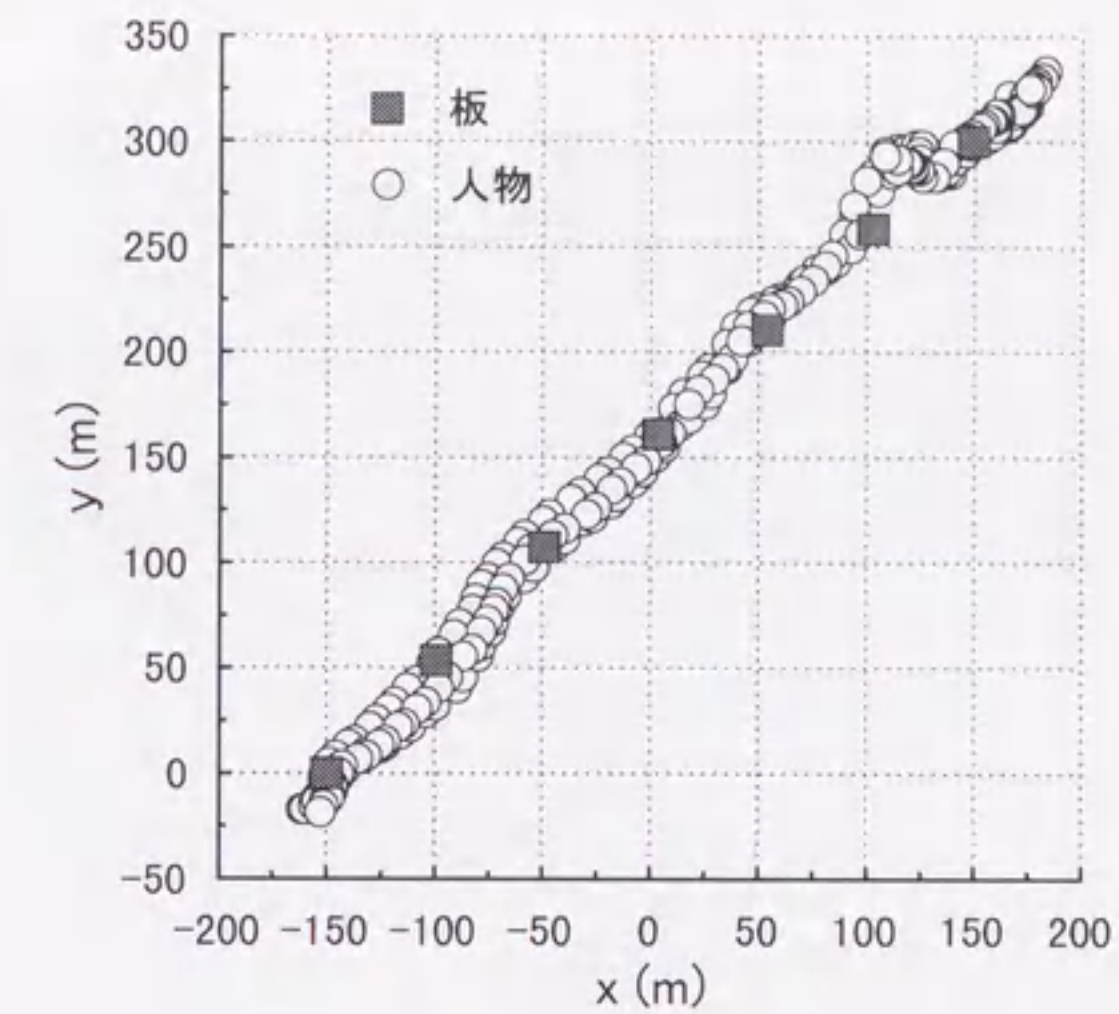


図 4.5: 位置検出の評価.

## 4.1.3 合成結果

図 4.6 と図 4.8 は、自然風景の中に人物が入り込んだ様子である。背景はテクスチャで、柵とテーブルの間は約  $1m$  である。

## i) 一人の場合

仮想空間に入り込む人物が一人の時は、図 3.2 のように、人物の存在する場所でカメラの光軸に対して垂直な面をクリッププレーンとする。この方法では、仮想空間内を自由に人が歩き回ることができる。図 4.6 の合成例は、背景テクスチャと柵の間に人物がいる場面(図 4.7)である。

## ii) 二人の場合

仮想空間に入り込む人物が二人の時は、図 3.4 のように、二人の人物を通る平面をクリッププレーンとする。この方法では仮想空間に二人入り込むことができるが、カメラ映像に





図 4.6: 仮想空間に一人の場合.

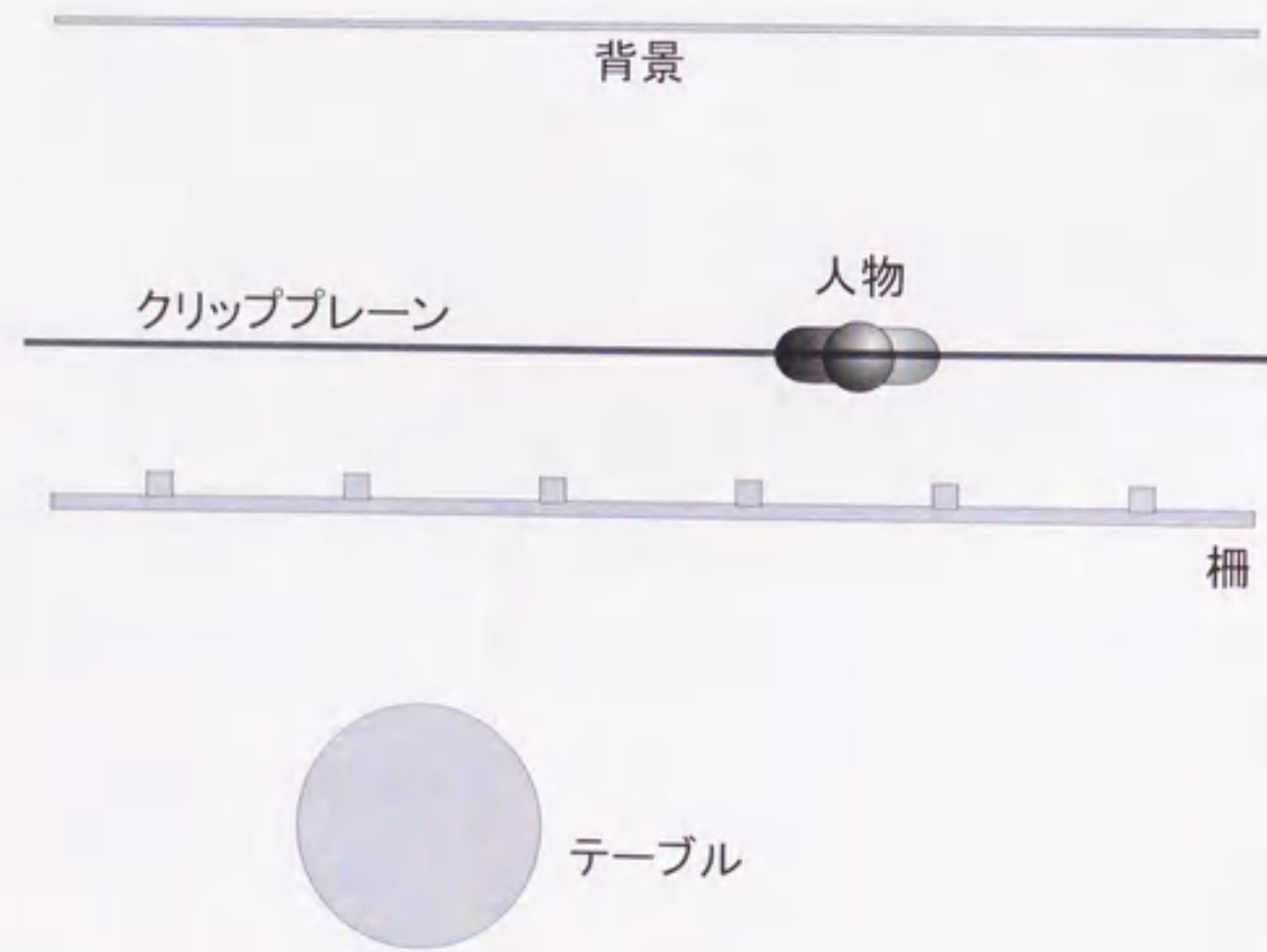


図 4.7: 仮想空間に一人の場合の配置図.

映っている二人に重なる部分があるときは、その重なり部分では仮想空間との前後関係を正しく表現することができない。図4.8の合成例は、図4.6の場面にさらに柵とテーブルの間に二人目が入り込んだ場面(図4.9)である。



図 4.8: 仮想空間に二人の場合.

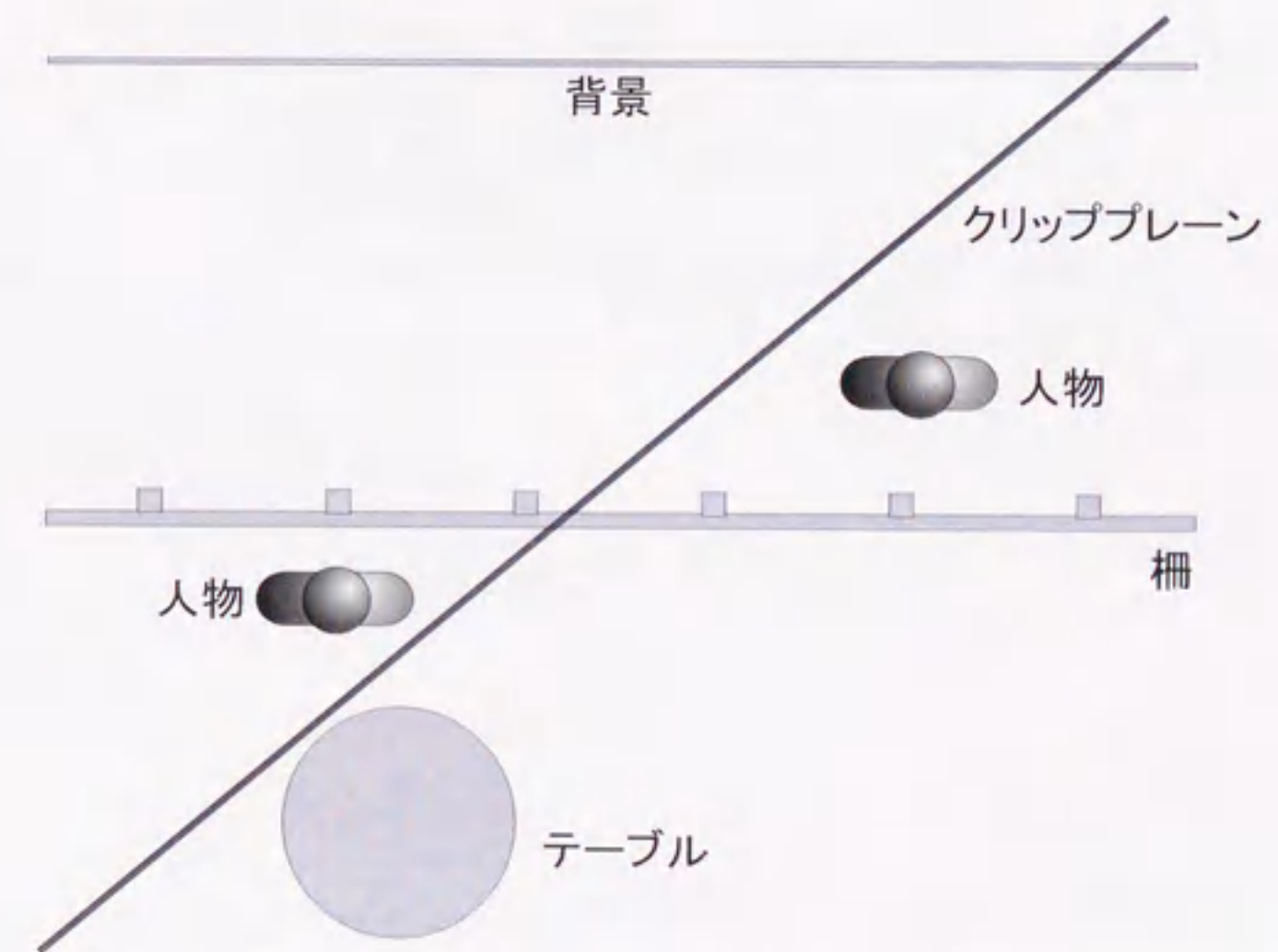


図 4.9: 仮想空間に二人の場合の配置図.



この場合、書割映像のレイヤー合成では5層必要で、さらに異なる層の二人の人物は別空間で撮影する必要がある。提案方法では必要な層は3層で、二人の人物を同じ空間で撮影できる。しかし、カメラ映像に映っている二人に重なる部分があるときは、その重なり部分では仮想空間との前後関係を正しく表現することができない。

## 4.2 連続した物体の合成例: 手を用いた組立作業

Mixed Reality の応用の一つとして論じられている分野に、組立作業などのトレーナや作業教示がある。本論文では、仮想空間のPCに現実の拡張ボードを取り付けるという状況を設定し、比較的狭い空間に手やボードといった相対的に大きな実物体を融合することを試みた。この実験のシステム構成を図4.10に示す。

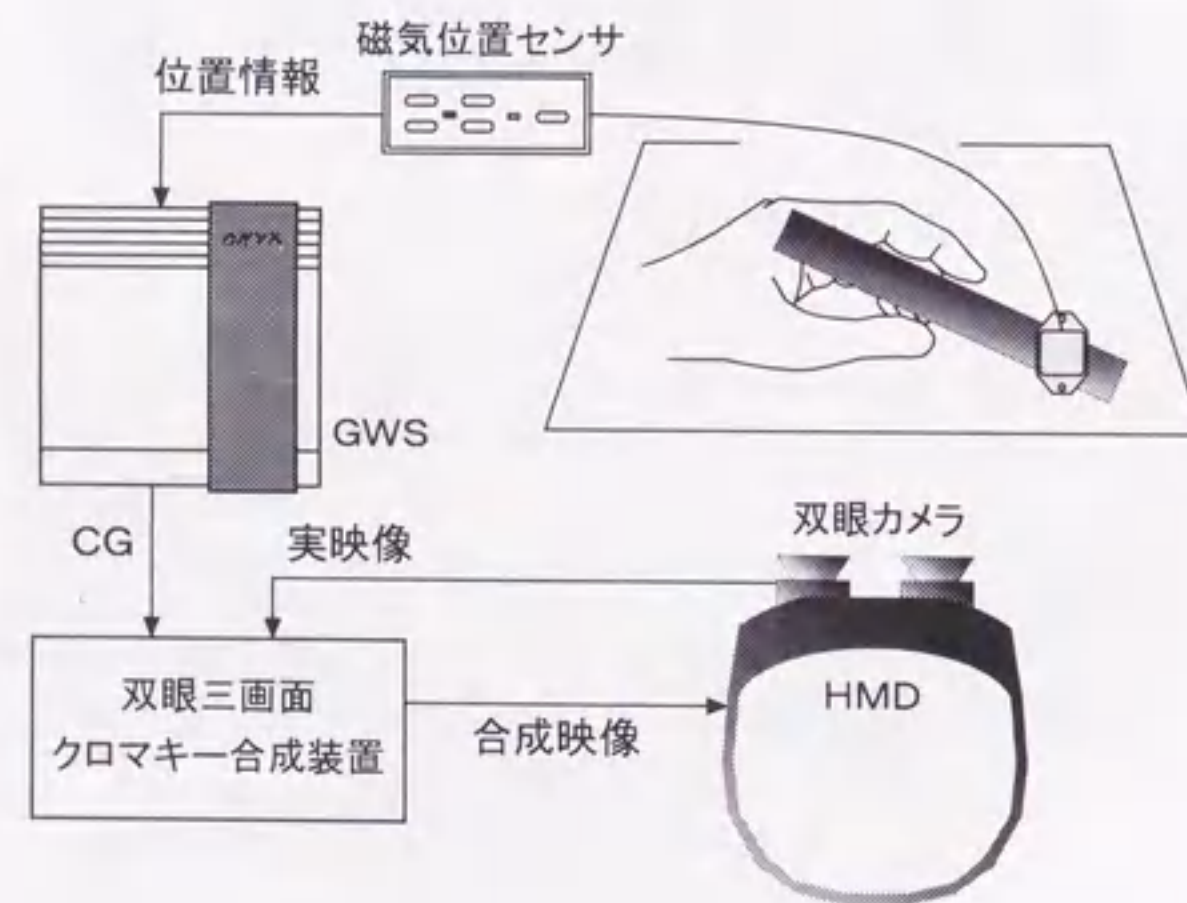


図 4.10: 手作業空間での実験。

### 4.2.1 磁気位置センサによる位置検出

実空間の物体の位置検出に、磁気位置センサ FASTRAK (表4.1) を用いた。磁気センサを拡張ボードに取り付けボードの位置・姿勢を検出し、ボードの面をクリッププレーンと

した。

表 4.1: FASTRAK の仕様

項目	仕様
メーカー	Polhemus Incorporated
計測方式	磁界ベクトル検出
計測範囲	最大 305cm, 全方向 (360°)
静的精度	0.08cmRMS, 0.15°RMS
分解能	1cm あたり 0.0002cm, 0.025°
遅延時間	4.0m 秒 (レシーバ測定サイクルの中央から出力ポートの転送開始まで)
サンプリング周波数	最大 120Hz

### 4.2.2 ビデオシースルー型 HMD を用いた結果提示

ビデオシースルー型 HMD には、視線上の眼に近い位置に小型カメラが取り付けられている。この小型カメラの映像を実映像としてクロマキー合成装置に入力し、合成結果を HMD に提示することによって、眼前の空間に仮想空間が融合しているかのような効果を得ることができる。なお、本実験では合成実験を主眼としているので、HMD の位置・姿勢は既知とし、固定して使用している。

### 4.2.3 合成結果

合成例を図4.11に示す。実空間の映像は手と手に持っている拡張ボードで、その他はすべて仮想空間の映像である。仮想空間のPCケースの内側に実空間の映像が入り込んでいるのが分かる。

書割映像のレイヤー合成システムでは、PC ケースのように連続的に囲まれた仮想物体



の内部に、実物体が自由に入り込むことはできないが、提案方法により可能になった。



図 4.11: PC の組み立て作業.

## 第5章 結論

本研究では、実空間と仮想空間との視覚的融合に問題となる隠蔽関係の表現方法について検討を行った。

まず、現在の合成システムの隠蔽関係表現の実現方法を考察した。そして現在の合成システムは核となる技術によって「単合成」「ピクセル合成」「レイヤー合成」として分類できることを示し、それぞれの長所・短所を明らかにした。その結果、ピクセル合成システムは原理的に正確な隠蔽関係を表現できるが、映像品質は研究段階であり、レイヤー合成システムは広く用いられており映像品質の点で有利であるが、隠蔽関係の表現力が乏しいことが明らかになった。そして、広く使用されており解像度の点で有利なクロマキー合成装置を用いて、「レイヤー合成」でありながら従来の方法より「ピクセル合成」の表現力に近くなる手法を考えるのが妥当であることが分かった。

さらに、改良方法を考察するに当たって、まずレイヤー合成の原理を明確にし、層の数・向き、合成映像の提示方法によってどのように隠蔽関係の表現力が変わるかを示した。その結果、既存のシステムは垂直分割・静的な層・単眼視であり、新たに斜方分割・動的な層・立体視を導入することによって、既存のシステムでも表現力を改善できることを明らかにした。

次に、この改良方法を実現する方法を提案した。実証実験には、岐阜大学バーチャルシステム・ラボラトリーに設置されている双眼クロマキー合成装置を使用した。この装置は各眼クロマキー面を2面持っているので、3層のレイヤー合成システムである。従って、本実験は3層分割で行った。さらに、この双眼クロマキー合成装置を用いてシステム構築するにあたって、実空間と仮想空間のキャリブレーション作業を画像処理により自動化することで、省力化・高精度化を図った。このとき、より多くの測定点を用いてキャリブレーション



ンをするために、自動的に推論しながらキャリブレーションをする方法について述べた。

最後に、提案方法の有効性の検証実験として、(a)必要な層の数を削減の例と(b)連続した仮想物体との合成例を示した。

実験(a)では、バーチャルスタジオに提案方法を応用した。登場人物が二人の場合では、垂直に層を決定した場合5層必要であるが、提案方法では3層で実現できることを示した。そして、空間内を自由に人が移動できることにより、実物体の移動範囲が拡大していることを示した。

実験(b)では、仮想物体のPCケース内部に実物体の部品を取り付ける作業の教示システムに応用した。PCケースは部品に対して周りを連続的に囲む仮想物体である。これまでのレイヤー合成では隠蔽関係の整合を保つためには移動範囲は限られていたが、本提案方法によって部品がPCケースの内部を自由に移動することができるようになった。

これら一連の研究により、現在のレイヤー合成システムの隠蔽関係の表現力を増す方法を提案し、その有効性を示すことができた。

この成果は合成システムの高性能化やコスト削減につながるため、産業・医療・アミューズメントなどの分野でVirtual Realityの技術応用が促進されることを期待している。

## 目次

1.1	Autonomy, Interaction and Presence.	1
1.2	Virtual Continuum.	3
1.3	ピクセル合成.	5
1.4	レイヤー合成.	6
2.1	対象空間の例.	9
2.2	対象空間の分割.	10
2.3	各層の映像.	10
2.4	分割空間の合成.	10
2.5	垂直2層分割	11
2.6	連続体の垂直3層分割	12
2.7	斜面2層分割.	13
2.8	斜面3層分割.	14
2.9	斜面3層分割による垂直5層分割表現.	14
2.10	静的な層.	15
2.11	動的な層.	16
2.12	正常な隠蔽関係.	16
2.13	物体同士の干渉.	17
2.14	ピクセル合成とレイヤー合成の分類.	18
3.1	3層の生成.	20



3.2 一人の場合.	21
3.3 クリッププレーン.	22
3.4 二人の場合.	22
3.5 システム構成.	23
3.6 SDP-4用のGWS生成画面.	25
3.7 IrisGL内の座標変換の流れ.	29
3.8 オブジェクト座標系.	30
3.9 ワールド座標系.	31
3.10 視野座標系.	31
3.11 クリップ座標系.	32
3.12 正規座標系.	33
3.13 ウィンドウ座標系.	33
3.14 画面座標系.	34
3.15 視点位置の計算.	36
3.16 視点位置の計算.	36
3.17 ウィンドウ変換.	37
3.18 クリッププレーンの計算.	38
3.19 キャリブレーション用立方体(1辺20cm).	41
3.20 手動キャリブレーションの様子.	41
3.21 自動キャリブレーションのシステム構成.	42
3.22 画像処理の流れ.	43
3.23 RGB入力画像.	44
3.24 色相画像.	46
3.25 ヒストグラム.	46
3.26 ラベリング.	47
3.27 自動多点キャリブレーションの誤差と標準偏差.	49
3.28 左右映像切り替え装置.	49

4.1 スタジオを使った実験.	52
4.2 検出処理の流れ.	53
4.3 検出処理中の画像.	55
4.4 キャリブレーションの方法.	56
4.5 位置検出の評価.	57
4.6 仮想空間に一人の場合.	58
4.7 仮想空間に一人の場合の配置図.	58
4.8 仮想空間に二人の場合.	59
4.9 仮想空間に二人の場合の配置図.	59
4.10 手作業空間での実験.	60
4.11 PCの組み立て作業.	62



## 表 目 次

3.1 PRIMATTE-1 の仕様 . . . . .	26
3.2 PRIMATTE-2 の仕様 . . . . .	26
3.3 Onyx の仕様 . . . . .	26
3.4 SDP-4 の仕様 . . . . .	27
3.5 VGA60 の仕様 . . . . .	27
3.6 Indigo2 HighImpact の仕様 . . . . .	44
4.1 FASTRAK の仕様 . . . . .	61



## 参考文献

- [1] Myron W. Krueger:「Artificial Reality II」, Addison-Wesley Publishing Company, Inc. (1991).
- [2] 廣瀬 通孝:「バーチャル・リアリティ応用戦略」, オーム社 (1991).
- [3] 藤本 英雄:「人工現実感の展開」, コロナ社 (1994).
- [4] D. Zeltzer:「Autonomy, Interaction and Presence」, PRESENCE, Vol.1 No.1, pp127-132 (1992).
- [5] 廣瀬 通孝:「バーチャルリアリティ」, オーム社 (1994).
- [6] Thomas H. Massie and J. K. Salisbury:「The PHANTOM Haptic Interface: A Device for Probing Virtual Objects」, Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (1994).
- [7] I. Sutherland:「The Ultimate Display」, Proceedings of IFIP Congress, pp.506-508 (1965).
- [8] I. Sutherland:「A Head-Mounted Three Dimensional Display」, Proceedings of Fall Joint Computer Conference Vol.33, pp.757-764 (1968).
- [9] C. Cruz-Neira, D. J. Sandin and T. A. DeFanti:「Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE」, Computer Graphics(Proceedings of ACM SIGGRAPH'93), pp.135-142 (1993).



- [10] Gurminder Singh, Steven K Feinder, Daniel Thalmann:「VIRTUAL REALITY SOFTWARE & TECHNOLOGY Proceedings of the VRST '94 Conference」, World Scientific Publishing Co. Pte. Ltd. (1994).
- [11] 岩田 洋夫 他:「人工現実感生成技術とその応用」, サイエンス社 (1992).
- [12] P.Milgram 他:「A Taxonomy of Mixed Reality Visual Displays」, IEICE Vol.E77-D No.12, pp.1321-1329 (1994).
- [13] Ronald T. Azuma:「A Survey of Augmented Reality」, Presence Vol.6 No.4, pp.355-385, (1997).
- [14] Stephen R. Ellis 他:「Pictorial communication in virtual and real environments Second edition」, Taylor & Francis Ltd.(1991,1993).
- [15] Ryugo Kijima and Takeo Ojika:「Transition Between Virtual Environment and Workstation Environment with Projective Head Mounted Display」, Proceedings of Virtual Reality Annual International Symposium, pp130-137 (1997).
- [16] 「人工現実感から複合現実感へ」, O plus E Vol.20 No.3, pp.346-359 (1998).
- [17] 宮田 一乗:「画像とCGの融合-コンピュータによる映像編集からデジタル合成まで-」, 画像ラボ Vol.9 No.1, pp.38-41 (1998).
- [18] 佐藤 清秀 他:「複合現実感における現実空間と仮想空間の融合に関する検討」, 日本バーチャルリアリティ学会第二回大会論文集, pp.161-164 (1997).
- [19] 中村 学, 村田 浩之 他:「現実・仮想融合空間における視知覚の整合化」, Human Interface News and Report Vol.12 No.2, pp.151-156 (1997).
- [20] 佐藤 清秀 他:「現実空間と仮想空間の位置合わせ手法-ステレオカメラと3次元センサの組み合わせ-」, 画像の認識・理解シンポジウム論文集I, pp.17-112 (1998).

- [21] 木島 竜吾 他:「バーチャルスタジオにおける照光システムシステムの開発」, 日本バーチャルリアリティ学会第二回大会論文集, pp.177-180 (1997).
- [22] 松田 隆夫:「視知覚」, 培風館 (1995).
- [23] 鈴木 伸介 他:「視差情報と遮蔽関係の矛盾の空間認知への影響について」, 日本バーチャルリアリティ学会大会論文集 Vol.1, pp.19-22 (1996).
- [24] Steven K. Feiner 他:「Architectural Anatomy」, Presence Vol.4 No.3, pp.318-326 (1995).
- [25] T. Kanade 他:「Virtualized Reality: Being mobile in a visual scene」, ICAT/VRST'95, pp.133-142 (1995).
- [26] 朴 鍾一 他:「イメージ表現ルームにおける3次元映像合成法」, 信学技報 IE97-168, pp.55-60 (1998).
- [27] 桑島 茂純:「PCベース3眼ステレオ・ビジョン・システム」, 画像ラボ第9巻第12号, pp.64-69 (1998).
- [28] 「特集:3次元情報のセンシング技術」, O plus E Vol.20 No.11, pp.1251-1286 (1998).
- [29] 林 正樹:「電子大道具による映像制作」, 第6回ヒューマン・インターフェース・シンポジウム論文集, pp.479-486 (1990).
- [30] 山田 茂 他:「VGA方式による3Dクロマキー合成システムの開発とその応用」, 日本バーチャルリアリティ学会第二回大会論文集, pp.173-176 (1997).
- [31] 「立体クロマキー合成システム」, (株)フォトロン (1997).
- [32] David F.Rogers, J.Alan Adams:「Mathematical Elements for Computer Graphics Second Edition」, McGraw-Hill,Inc. (1990).
- [33] 「Graphics Library Programming Guide」, Silicon Graphics, Inc..



- [34] 安居院 猛：「画像の処理と認識」，昭晃堂 (1992).
- [35] 日本図学会：「CGハンドブック」，森北出版株式会社 (1989).
- [36] 高木 幹雄，下田 陽久：「画像解析ハンドブック」，東京大学出版会 (1991).
- [37] Eugene Leonard：「Consideration Regarding the Use of Digital Data to Generate Video Background」，SMPTE J Vol.87, pp.490-504 (1978).
- [38] 林 正樹 他：「仮想スタジオ ～バーチャルリアリティのデスクトップ映像制作への応用～」，8th Symposium on Human Interface, pp.449-453 (1992).
- [39] 榎並 和雅：「マルチメディア時代の番組作成技術 -デスクトッププログラムプロダクション DTPP-」，NHK 技研 R&D No.42 (1996).
- [40] Pattie Maes 他：「The ALIVE System: Full-body Interaction with Autonomous Agents」，Proceedings of Computer Animation'95 Conference, pp.11-18, (1995).
- [41] Michiko Kusahara, Christa Sommerer and Laurent Mignonneau：「Art as Living System」，システム/制御/情報, pp.344-351 (1996).

## 謝辞

本論文をまとめるに当たり，終始ご親切なご指導を賜りました岐阜大学工学部応用情報学科 小鹿丈夫教授に心から感謝の意を表します。

また，有益なご討論，ご助言をいただきました岐阜大学工学部応用情報学科木島竜吾講師に感謝いたします。

最後に，本研究を遂行するにあたって，ご指導，ご協力を頂いた岐阜大学工学部応用情報学科小鹿・木島研究室の皆さま，岐阜大学バーチャルシステム・ラボラトリーの関係者に感謝申し上げます。



