

屋外環境における指示共有のための
複合現実型ジオラマインタフェース

林 将之

システム情報工学研究科
筑波大学

2013年 7月

概要

現代では、スマートフォンのような高機能携帯電話の普及により、遠隔地にいる人々の間で、音声やテキスト、画像、地図上の位置情報を用いたコミュニケーションが可能となっている。しかし、人々が共同作業を行う際に、注目して欲しい場所の位置情報を共同作業者に伝えるなど、空間中の特定の位置への指示を行う場面においては、指差しやレーザポインタなどが利用されている。

本論文では、スマートフォンやタブレットなどのカメラ付きのモバイル端末上に、周辺環境のミニチュア CG モデル（ジオラマモデル）を提示し、そのジオラマモデル上に矢印を提示することによって的確な 3 次元位置情報の指示を実現する、複合現実（Mixed-Reality; MR）型ジオラマインタフェースを提案する。これは、モバイル端末のカメラで撮影した実写映像上にジオラマモデルを重畳表示し、ユーザにモバイル端末のディスプレイをのぞき窓として、目の前のジオラマモデルを観察しているかのような映像を提示するインタフェースである。

提案インタフェースはジオラマモデルを介した間接的な指示方法であるため、ユーザは指示位置を理解するために、ジオラマモデルと現実空間との対応を把握する必要がある。本論文では、現実空間との対応を把握しやすいジオラマモデルの提示を行うことで、指示位置がわかりやすい複合現実型ジオラマインタフェースの設計を被験者実験により追求する。

被験者実験では、まず、ジオラマモデルを実写映像へ重畳することの是非、および現実空間の建物と同じ向きにジオラマモデルを整列させる方式の有効性を検証する。次に、ジオラマモデルの整列性を維持する仮想視点の操作方式を 3 種類挙げ、指示位置のわかりやすさと仮想視点操作のしやすさを比較する。これらの実験から、指示位置がわかりやすく、仮想視点操作がしやすい複合現実型ジオラマインタフェースの設計を行う。さらに、提案インタフェースの有効性を評価するため、現実空間中にアノテーションを直接重畳する拡張現実（Augmented-Reality; AR）型のインタフェースとの比較を行う。この比較実験では、指示共有を必要とする屋外作業の例として、屋外カメラ設置におけるポーズ決定タスクを用いる。

これらの被験者実験を通して、屋外における指示共有の新しい実現方法として、複合現実型ジオラマインタフェースの有効性を明らかにする。

目次

第 1 章	序論	1
第 2 章	3 次元位置情報の指示共有	4
2.1	直接的指示と間接的指示	4
2.2	地図を用いた指示共有	5
2.3	拡張現実感／複合現実感技術を用いた指示共有	6
2.4	World in Miniature	7
2.5	ジオラマモデルの構築	8
第 3 章	複合現実型ジオラマインタフェース	9
3.1	コンセプト	9
3.2	指示位置の入力	11
3.3	指示位置の提示	12
3.4	実写映像への重畳表示	14
3.4.1	実験環境	15
3.4.2	ユーザタスク	16
3.4.3	実験手順	16
3.4.4	実験結果	17
3.4.5	考察	20
第 4 章	複合現実型ジオラマインタフェース による指示共有	21
4.1	遠隔指示共有	21
4.2	画面の構成	22
4.3	音声指示の併用	23
第 5 章	ジオラマモデルの視点操作	24
5.1	整列効果	24
5.1.1	整列／非整列方式の比較	25
5.1.2	実験結果と考察	26
5.2	整列性を維持した視点操作方式	27
5.2.1	Augmented-Reality (AR)方式	27
5.2.2	Pan/tilt 方式	29
5.2.3	Dolly-round/crane (D/C)方式	30
5.3	整列性を維持した視点操作方式の比較実験	31
5.3.1	実験環境	31
5.3.2	ユーザタスク	32

5.3.3	実験手順	32
5.3.4	実験結果	32
5.3.5	考察	34
5.4	Pan/tilt 方式と Dolly-round/crane 方式の比較実験	35
5.4.1	A 地点の実験結果	36
5.4.2	B 地点の実験結果	38
5.4.3	考察	40
第 6 章	直接的な指示共有方式との比較実験	42
6.1	AR アノテーションインタフェース	42
6.2	撮影範囲の可視化	44
6.3	モバイル端末の位置姿勢推定	47
6.3.1	ビジョンベースの位置姿勢推定手法	47
6.3.2	マップの事前構築手法による PTAMM の初期化处理	47
6.4	屋外カメラ設置におけるポーズ決定タスク	48
6.4.1	タスクの概要	48
6.4.2	屋外カメラ設置におけるポーズ決定タスク用インタフェース	51
6.4.3	実験手順	51
6.5	実験結果と考察	52
6.5.1	印象評価の結果と考察	52
6.5.2	タスク所要時間の結果と考察	54
第 7 章	結論	56
	謝辞	59
付録 A.	PTAM の精度評価	66
	3 次元点群の投影誤差による評価	67
	評価手順	68
	世界座標系の変換	69
	評価結果	69
付録 B.	性能諸元	72
	モバイルカメラ(Creative 製 VF0490)仕様	72
	慣性センサ(INTERSENSE 製 InertiaCube3)仕様	72
	ユーザ端末(SONY 製 VGN-UX90PS)仕様	73
	ユーザ端末カメラ(SONY 製 VGN-UX90PS 背面内臓カメラ)仕様	73
	ユーザ端末(Acer 製 ICONIA TAB W500S)仕様	74
	ユーザ端末カメラ(IDS 製 UI-1221LE-C-HQ)仕様	74
	公表論文リスト	75

図目次

図 1-1	指差しによる指示（左）と建築模型を用いた指示（右）の例.....	1
図 1-2	複合現実型ジオラマインタフェースのコンセプト	2
図 3-1	タッチスクリーンからのレイキャストを用いた指示位置の入力	11
図 3-2	指示共有の実現.....	13
図 3-3	ジオラマモデル上の矢印による指示位置の提示例.....	13
図 3-4	実写映像への重畳表示の有無.....	14
図 3-5	実験環境の航空写真（左）と実験に用いたジオラマモデル（右）	15
図 3-6	A 地点で行う写真撮影タスクに用いる指示位置.....	16
図 3-7	実写映像への重畳表示の有無の比較実験 印象評価結果	18
図 3-8	実写映像への重畳表示の有無の比較 タスク所要時間.....	19
図 4-1	無線ネットワークを介した提案インタフェースによる指示共有方式.....	22
図 4-2	指示共有システムの画面構成.....	23
図 5-1	整列方式と非整列方式.....	24
図 5-2	実験環境（C 地点）の航空写真（左）と実験に用いたジオラマモデル（右）	25
図 5-3	整列方式／非整列方式の比較 タスク所要時間.....	27
図 5-4	Augmented-Reality (AR)方式	28
図 5-5	視点の回りこみ操作例.....	29
図 5-6	Pan/tilt 方式	30
図 5-7	Dolly-round/crane (D/C)方式.....	31
図 5-8	整列性を維持した視点操作方式の比較 主観的な印象評価	33
図 5-9	整列性を維持した視点操作方式の比較 タスクの所要時間	34
図 5-10	Pan/tilt 方式と D/C 方式の比較実験 A 地点 印象評価結果.....	36
図 5-11	Pan/tilt 方式と D/C 方式の比較実験 A 地点 タスク所要時間	37
図 5-12	B 地点で行う写真撮影タスクに用いる指示位置.....	38
図 5-13	Pan/tilt 方式と D/C 方式の比較実験 B 地点 印象評価結果.....	39
図 5-14	Pan/tilt 方式と D/C 方式の比較実験 B 地点 タスク所要時間	40
図 6-1	AR アノテーションインタフェース	43
図 6-2	撮影範囲を可視化するための視体積の表示方式.....	45
図 6-3	ジオラマモデル上でカメラの撮影範囲を表示した場合のわかりやすさ	46
図 6-4	AR アノテーション型でカメラの撮影範囲を表示した場合のわかりやすさ	46
図 6-5	マップの事前構築による PTAMM の初期化	48
図 6-6	実験環境（左）とジオラマモデル（右）	50

図 6-7	屋外カメラ設置禁止場所を示す×マーク	50
図 6-8	屋外カメラ設置におけるポーズ決定タスクのための画面構成	51
図 6-9	主観的な印象評価の結果	53
図 6-10	第 6 章 直接的な指示共有方式との比較実験 被験者毎の所要時間	54
図 6-11	直接的な指示共有方式との比較実験 タスク毎の平均所要時間	55
図 7-1	3次元点群の配置と, その投影誤差による位置姿勢の推定誤差評価	68
図 7-2	NAIST Campus Package 02 のサンプル画像	70
図 7-3	PTAM の位置姿勢推定の誤差	71
図 7-4	3次元点群の投影誤差	71

表目次

表 3-1	主観的な印象に関する質問	17
表 3-2	実写映像への重畳表示の有無の比較 タスク所要時間の平均.....	18
表 3-3	実写映像への重畳表示の有無の比較 タッチパネル操作の利用状況.....	19
表 5-1	整列方式／非整列方式の比較 主観的な印象に関する質問	26
表 5-2	写真撮影タスク所要時間の平均	26
表 5-3	整列性を維持した視点操作方式の比較 タスク所要時間の平均.....	34
表 5-4	Pan/tilt 方式と Dolly-round/crane 方式の比較実験 タスク所要時間の平均.....	36
表 5-5	Pan/tilt 方式と D/C 方式の比較実験 A 地点 タッチパネル操作の利用状況	37
表 5-6	Pan/tilt 方式と D/C 方式の比較実験 B 地点 タッチパネル操作の利用状況	40
表 6-1	直接的な指示共有方式との比較実験 主観的な印象に関する質問	53

第1章 序論

現代では、スマートフォンのような無線ネットワークを介するモバイル端末の普及により、遠隔地にいる人々の間で、音声やテキスト、画像、地図上の位置情報などを用いたコミュニケーションが可能となっている。しかし、人々が共同作業を行う際に注目して欲しい場所を共同作業者に伝えるなど、空間中の特定の位置への指示を他者と共有する場面（これを指示共有と呼ぶ）では、指差し、指し棒、レーザポインタなどを用いて、指示したい位置を直接指し示すという単純な方法が用いられている（図 1-1 左）。こうした指示共有方法は、卓上や室内空間においては、空間中の特定の位置を数十センチ～数センチ以下の高い精度で簡単に指示できる一方で、指示位置が見える場所に人々が集合する必要があることや、指示位置が遠い場合に指示が曖昧となる（ポインタが見えにくくなる）ことから、屋外など広い作業環境においては常に便利な方法とは限らない。

建築の分野では、建築模型（建物や都市の実物の模型）を用意し、その各部を見ながら相談するという方法が従来から用いられている [1] [2]。建築模型を用いることで、人々が建物の周辺や都市の中を移動して指示位置の近くに行かなくても指示共有が可能になる（図 1-1 右）。しかし、人々が建築模型の周囲に集合しなければならないという不便が残る。また、建築模型は持ち運びに不向きなため、屋外で気軽に使うことができない。

屋外環境で共同作業を行う場合に、環境中に展開している人々が、各人の持ち場に居ながら、互いに的確な 3 次元位置情報の指示共有を行える効果的な方法は、未だ実用化されていない。



図 1-1 指差しによる指示（左）と建築模型を用いた指示（右）の例

本論文では、スマートフォンやタブレットなどのカメラ付きのモバイル端末上に、周辺環境のミニチュア CG モデル（ジオラマモデル）を提示し、そのジオラマモデルを介して指示共有を実現する、複合現実（Mixed-Reality; MR）型のジオラマインタフェースを提案する。提案インタフェースは、図 1-2 に示すように、ユーザに対し、モバイル端末の画面を通して、ジオラマモデルがあたかもユーザの目前に存在しているかのような映像を提示する。ユーザがタッチスクリーンを用いてジオラマモデル上の任意の点を入力すると、そこに指示位置を示す矢印（アイコン）が表示される。

さらに、提案インタフェースによって入力された指示位置の情報を、無線ネットワークを介してユーザ同士で共有することにより、各ユーザが屋外環境中のどこに居ても指示共有できる手法を実現する。

提案インタフェースはジオラマモデルを介した間接的な指示方法であるといえる。ユーザは、ジオラマモデル上に表示された矢印を見て、指示位置が現実空間中のどこであるかを理解するために、ジオラマモデルと現実空間との対応を把握する必要がある。このためには、複合現実型の提示により視覚的な手がかりを与えることが有効と考えられる。本論文では、指示位置を理解しやすい複合現実型ジオラマインタフェースの設計について被験者実験を通して追求する。さらに、拡張現実感（Augmented-Reality; AR）技術を用いた直接的なアノテーション提示方法との比較を行い、屋外環境における複合現実型ジオラマインタフェースの有効性を評価する。

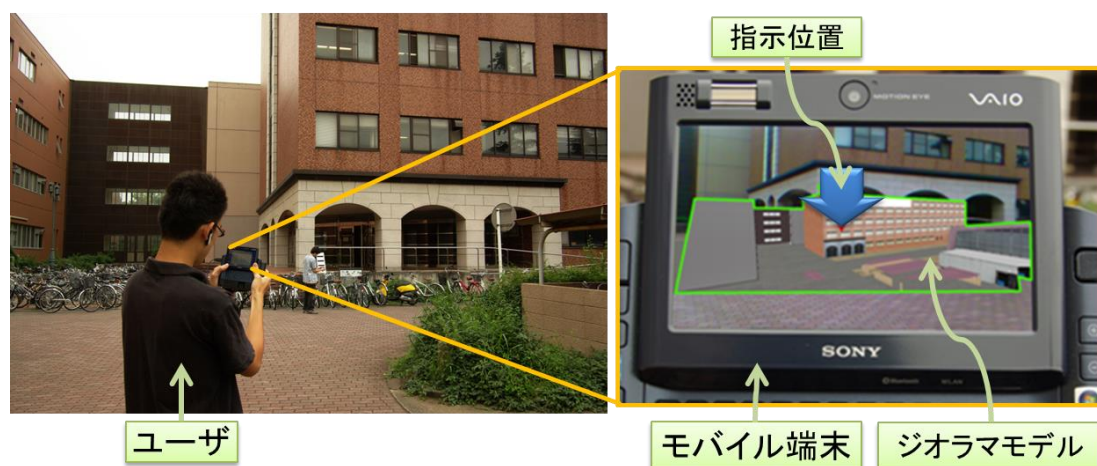


図 1-2 複合現実型ジオラマインタフェースのコンセプト

以下、本論文では、第2章で3次元位置情報の指示共有を実現する従来手法と、本論文のアプローチについて述べる。第3章では、複合現実型ジオラマインタフェースのコンセプトや設計について述べる。このインタフェースは、3次元位置情報の入力・提示を行うユーザインタフェースであるため、具体的にどのようにして遠隔地の人物と指示共有を実現するのかについて、第4章で述べる。第5章では、モバイル端末を用いたジオラマモデルの仮想視点を操作する方式について、被験者実験に基づいた比較検討を行う。第6章では、AR 技術を用いて現実空間中の物体

に直接的なアノテーションを提示する指示共有方法との比較を行うための被験者実験について述べる．ここでは，屋外作業の例として屋外定点カメラの設置相談タスクを挙げる．そして最後に第 7 章で本論文の結論を述べる．

第2章 3 次元位置情報の指示共有

本章では、位置情報の指示共有に関連する既存研究を取り上げ、本論文との差異を明らかにする。

2.1 直接的指示と間接的指示

本論文では、指差しや指し棒、レーザポインタなどを用いて、指示したい位置や物体を現実空間または実写映像上で直接指し示す指示共有方法を、直接的指示と呼ぶ。直接的指示には、拡張現実感 (Augmented-Reality; AR) 技術 [3]を用いて現実空間の物体に矢印などのアノテーションを提示する方式 [4] [5] [6] [7] [8]も含む。

指差し／レーザポインタなどの単純な直接的指示には以下に示す制約があり、特に屋外のような広い空間で用いる場合には大きな制限となっている。

- (i) 指示位置が見える場所に人々が集合する必要がある。
- (ii) 人々と指示位置との距離が離れるほど、指示の精度が低下する。

制約(i)の技術的な解決を目指した直接的指示方式としては、遠隔会議における指差しの支援技術 [9] [10] [11]や遠隔操作可能な装着型レーザポインタ [12]が挙げられる。これらの方法では、指示者が遠隔地に居る被指示者に指示を提示することができる。しかし、これらの手法においても、被指示者から近い場所を指示された場合にはわかりやすいが、被指示者から遠い場所を指示された場合には、指示が曖昧になりやすいという制約(ii)が残る。そのため、被指示者は指示位置・指示対象物の近くにいない必要がある。制約(ii)が生じる原因として、指や指し棒が届かない、レーザポインタが見えにくい、遠くの物体は見た目が小さい、などの理由が挙げられる。

一方、地図 [13]や写真 [14] [15]、言葉による説明などを用いて、間接的に指示位置を示す指示共有方法を、間接的指示と呼ぶ。これらの指示共有方法は、スマートフォンやタブレット端末を用いることで、遠隔地にいる人々の間でも指示共有が可能である。また、間接的指示は、ユーザから離れた場所でも正確に指示できるため、直接的指示における制約(i), (ii)は存在しない。

間接的指示方式として、目的地までのルートナビゲーションサービスのように、地図に基づいた移動経路指示は既に実用化されているが、モバイル端末の利用者同士が3次元の位置情報を的確に指示共有するための方法は、いまだ未開拓の研究分野である。

2.2 地図を用いた指示共有

一般的に、空間知識は、場所の見た目に関するランドマーク知識、道筋に関するルート知識、それらを統合して環境の全体像を構築するサーベイ知識の3形式に分けられると言われている[16]。写真を用いた指示や直接的指示の場合は、ランドマーク情報のみを伝達するが、地図を用いることで、ルート知識とサーベイ知識を与えることができる。

ランドマーク情報に基づく空間内のナビゲーションは、人間にとって最も自然なナビゲーション手法であると言われている。写真を用いた指示や直接的指示では、視覚的なランドマーク情報を提示する。一方、2次元の地図は抽象的なランドマーク情報を提示する。2次元地図を用いた指示共有の例として、Nilsson ら [13]が行った地図上に作業者の配置などを重畳表示するユーザインタフェースに関する、優れた評価実験がある。このシステムは、2次元地図のまわりにいる複数の責任者が現場にいる作業者に対して与える指示を相談するために設計されており、現場にいる作業者がシステムを携帯することが難しい。そのため、作業者同士が相互に指示を与えるという用途には向いていない。

Fujita ら [17]は、2次元地図を用いたナビゲーションにおいて、視点の平行移動操作を拡大／縮小操作やチルト操作にも対応付けたインタフェースを提案した。彼らは、ペン入力に対応した画面を用いて2次元地図を観察する際の操作方法に関する評価を行っており、室内での利用が前提となっている。

近年では、建物の3次元形状と写実的なテクスチャを持つモデル（ジオラマモデル）を航空写真などから都市レベルで復元し、モバイル端末へ配信する技術が普及しつつあり、3次元地図の利用が現実的になっている。ジオラマモデルは、写実的な3次元地図であり、これを様々な視点から観察することにより、サーベイ知識を含む全ての形式の空間知識を得ることができる。

本論文で提案する複合現実型ジオラマインタフェースは、MR 技術を用いてジオラマモデルを提示し、そのジオラマモデル上で遠隔地に居る人々と位置情報のコミュニケーションを行うものである。本論文では、モバイル端末を携帯するユーザが屋外環境中におり、実際の風景を見ている状態で、指示位置がわかりやすいと感じられるジオラマモデルの提示方式の追求に取り組む。

ジオラマモデルを用いて位置情報を提示する技術は、大規模な仮想空間内のナビゲーションにおける重要な要素技術であるため、ヒューマンファクタを考慮した様々な研究がなされてきた。Nurminen ら [18]は、写実的な3次元地図は2次元地図に比べ、環境を認知しやすくする視覚情報を与える一方で、効果的なナビゲーションを実現するためには、モバイル端末を用いたインタラクションの適切な設計が必要であると述べている。本論文の取り組みは、モバイル端末向けの複合現実感技術を用いて、ジオラマモデルとのよりよいインタラクションを求める研究の一つであるといえる。

2.3 拡張現実感／複合現実感技術を用いた指示共有

拡張現実感 (Augmented-Reality; AR) 技術は、光学シースルーまたはビデオシースルーのディスプレイを用いて、現実空間の風景に、仮想空間の CG 映像を重畳表示する技術である。一方、CG で構築されたジオラマモデルのように、現実空間から取り込んだ情報を用いて仮想空間を構築する技術は、拡張仮想感 (Augmented-Virtuality; AV) と呼ばれている。複合現実感 (Mixed-Reality; MR) は、AR 技術や AV 技術のように現実空間と仮想空間の情報を融合して提示する技術全般を指す言葉として定義されている [3]。

AR 技術を用いた位置情報の入力・提示方法として、Höllerer ら [4]は、屋外歩行者向けに光学シースルー型 HMD (Head Mounted Display)を用いて実際の風景にアノテーションを表示するインタフェースや、モバイル端末上に地図を表示するインタフェース、室内ユーザ向けに屋外環境の CG ジオラマモデルを AR 提示するインタフェースなど、状況に合わせた複数の AR 型のインタフェースを提案している。彼らのコンセプトは、ユーザの状況／役割に応じて適切なインタフェースを利用し、位置情報は全てのインタフェースで共有するというものである。

本論文で提案するインタフェースは、写真から取り込んだテクスチャを持つ写実的なジオラマモデルを、実写映像上へ AR 表示する複合現実型のインタフェースである。これは Höllerer らが HMD を用いて屋内ユーザ向けに提案した、ジオラマモデルをテーブルの上に AR 提示するインタフェースのアイデアと類似している。

本論文の提案インタフェースは、屋外での利用を前提とし、モバイル端末を用いて実現する。屋外において、写実的なジオラマモデルを提示することで、現実空間とジオラマモデルとの対応関係を把握するための、視覚的な手がかりを与える。そのため、ジオラマモデルの座標系で位置情報を提示すれば、ユーザは、その位置が現実空間中のどこに相当するのかを理解することができる。この方法は、光学シースルー／ビデオシースルーのディスプレイを用いて現実空間の風景にアノテーション（物体や位置を指し示す注釈情報）を直接重畳する手法と比べて、CG と現実空間の位置合わせを正確に行う必要がないという利点があると考えている。

AR 技術を用いたアノテーション表示が空間中の位置情報の提示に有効であることは、これまで様々な屋内作業やその評価実験を通して明らかになっている [19] [20]。しかし、屋外における AR 技術を用いたアノテーション表示には、奥行き（視点位置からアノテーションまでの距離）を正しく知覚することが難しい [5]、遮蔽物による隠れが生じる、遠い場所は見た目が小さくなるので細かい部分が見えにくい、という問題点がある。

奥行きの知覚精度を改善する映像表現について過去に様々な評価が行われた [21]。その結果、奥行きの知覚精度を改善する映像表現はいくつか提案された [4] [6]。しかし、問題の根本的な解決には至っていない。

屋外環境では、屋内環境に比べ、指示された場所へユーザが移動して視認するのに要する時間や手間が増える。移動の必要が生じる状況として、指示された位置が遠い場合や、遮蔽物により

指示位置が見通せない場合が考えられる。シースルービジョン [22] [23] [24]や注目物体の強調表示 [25]の利用により、遮蔽物と指示位置の隠蔽関係を変え、指示位置を仮想的に見ることもできるが、アノテーションがユーザから見て建物の裏側となる壁を指している場合などにはその指示位置を理解するのが困難である。

一方、ジオラマモデルのような3次元地図上に表示した指示位置により指示共有を行う場合、現実空間とCGの位置合わせに誤差が生じても指示位置の示す位置は変わらないため、高精度の位置合わせが必要という技術的な問題は回避される。さらに、ジオラマモデルを観察する仮想視点の移動や縮尺の変更により、距離の制約からも解放される。ただし、仮想視点と現実空間におけるユーザの視点が大きく異なる場合には、2次元地図よりは容易であるものの、ジオラマモデルと現実空間の対応関係を把握することがユーザにとって困難になりうることに注意する必要がある。

提案インタフェースは、屋外環境とそのジオラマモデルとを同時に見ることができるとする状況を前提としており、指示位置がユーザの視野内にある場合に最も効果を発揮する。一方、Sandor ら [26]は、3次元空間を歪曲させる手法により、ユーザの視野外にある指示位置を可視化した。特に、Radial Distort と呼ばれる手法は、建物の裏側となる壁も可視化できる。このような指示位置の提示手法は、提案インタフェースと組み合わせることが可能であり、それぞれの手法のユーザビリティを明らかにすることで、より優れた指示共有インタフェースの検討が可能となる。

2.4 World in Miniature

Head Mounted Display (HMD)を用いて World in Miniature (WIM) と呼ばれる周辺環境の3次元モデルをユーザの目前に提示するコンセプトは、仮想環境のナビゲーション／インタラクションを支援する技術として Stoakley ら [27]によって提案された。これは、室内空間での利用を想定した Virtual-Reality システムで、仮想物体の配置を操作することを目的としていた。Wingrave ら [28]は都市レベルの広域環境での利用を想定した Scaled and Scrolling WIM を提案し、WIM のスケール変更や平行移動を行いながら環境を探索する際のユーザビリティについて、被験者実験により定量的、定性的側面から詳細に評価した。

現実空間の WIM 表現を用いて、AR 型の歩行者ナビゲーションへ応用した初期の取り組みに、Höllerer ら [29]のビデオシースルーHMDを用いたものがある。Höllerer らが用いた WIM は建物内部の通路や部屋をワイヤフレームで表現したモデルであり、壁や床の模様のように、ランドマーク知識となる情報は含まれていなかったものの、カメラの位置姿勢推定精度が低い場合でも、WIM の提示はナビゲーションに有効であることが示された。

近年では、大隈ら [30]が科学ミュージアムの建物内で3次元地図を使った展示物へのナビゲーションを行う方式を提案し、被験者実験により評価した。彼らのナビゲーション方式は、3次元地図を見下ろす視点を用いる点と、ユーザの向きに合わせて視点を自動回転させることで整列効

果を利用している点で提案インタフェースと類似している。しかし、本論文では、指示位置を理解しやすくするために、屋外環境との対応が把握しやすいジオラマモデルの操作・提示方式を追求しており、大隈らの実験とは評価対象と状況が異なっている。

2.5 ジオラマモデルの構築

本論文で提案する複合現実型ジオラマインタフェースを実現するためには、ジオラマモデルとして屋外環境の写実的な 3 次元 CG モデルを予め構築しておく必要がある。写実的な CG モデルが利用できるならば、提案インタフェースは GPS (Global Positioning System) と 3 軸のジャイロセンサ (慣性センサ)、加速度センサ、電子コンパスなど、既に多数のスマートフォンやタブレット PC に搭載されているセンサを用いたモバイル端末の位置姿勢によって実現できる。

都市全体の写実的な CG モデル (3 次元地図) を構築することは、膨大な手間がかかる。そのため、携帯電話と 3G ネットワークを用いて 3 次元地図をモバイル端末へ配信・描画する技術 [31] が実現されながらも、都市全体の 3 次元地図は誰もが容易に入手できるものではなかった。しかし、近年では、航空写真や LiDAR、インターネット上で公開されている多数の写真などを用いた自動モデル化技術の発展 [32] [33] によって、都市全体を効率的にモデル化することが可能になってきている。それに伴い、Google 社、Apple 社、Nokia 社などがスマートフォンやタブレット向けに、一部の都市の写実的な CG モデル (3 次元地図) のオンライン配信サービスを開始しており、3 次元地図の利用が一般に普及しつつある。本論文で提案する複合現実型ジオラマインタフェースは、このような 3 次元地図配信サービスとスマートフォン／タブレットを組み合わせることで実現可能であり、屋外環境において 3 次元位置情報の指示共有を行う新たな手法を提供することができる。

第3章 複合現実型ジオラマインタフェース

屋外環境において的確な3次元位置情報の指示共有を行うために、本論文では複合現実型ジオラマインタフェースを提案する。これは、予め作成されたユーザの周辺環境の3次元CGモデル（ジオラマモデル）を用いて指示位置の入力と提示を行うインタフェースである。複合現実感技術を用いて、ジオラマモデルがあたかもユーザの目前に存在するかのように提示することで、実物の建築模型を観察する場合と同様に、ユーザの周辺環境を様々な視点から観察できる。本章では、このインタフェースのコンセプトの詳細と、指示位置の入力方法、指示位置の提示方法について述べる。

3.1 コンセプト

ジオラマモデルや建築模型のように、現実空間の3次元モデルを用いて位置情報を提示する場合、ユーザは現実空間とその3次元モデルとの対応を、視覚的な手がかりによって把握する必要がある[18]。ユーザは、実空間の風景と、仮想視点から見える仮想空間の風景とを比較し、ユーザ自身の現在地と仮想視点の位置関係、現実空間中での指示位置、を理解する必要がある。

本論文では、ランドマーク情報となる視覚的な手がかりを与えるために、建物の正確な形状と位置、写真ベースのテクスチャを持った、写実的なCGモデルをジオラマモデルとして用いる。しかし、このような写実的なジオラマモデルを用いても、ジオラマモデルを観察する仮想視点とユーザの視点と全く関連していない場合、ユーザは自身の記憶に頼って現実空間との対応を把握する必要がある。

本論文では、複合現実感技術を用いてユーザの実際の視点と関連した仮想視点を設定し、実際に見ている風景とジオラマモデルを比較できるようにすることで、ユーザ自身の現在地と仮想視点の位置関係を理解しやすくする。同時に、ジオラマモデル上のアイコン（矢印）で提示されている位置に対応する、実空間中での位置を理解しやすいインタフェースの実現を目指す。

屋外においてジオラマモデルを表示するためのディスプレイデバイスとしては、HMD (Head Mounted Display) や、スマートフォンやタブレットのようなモバイル端末が考えられる。利用するディスプレイデバイスによって、適したインタフェースの設計は異なると考えられている[4][34]。

HMDによる映像提示は、没入感が高い、ユーザの視点からのAR/MR映像を提示できる、ユーザは両手に何も持つ必要がないという利点がある。しかし、現在のHMDには、人間の視野よ

りも画面の視野が狭く、障害物などに注意しながら作業をすることが困難であることや、電源や映像送信用のケーブルによりユーザが動きにくくなるという問題がある。また、指示位置の入力をするためのインタフェースをどのように実現するかという問題も生じる。そのため、屋外での利用を想定する本論文では、モバイル端末を用いることを前提として、提案インタフェースの提示・操作方式を設計する。

本論文で提案する複合現実型ジオラマインタフェースには、AR 技術を用いて現実空間の風景にアノテーションを提示する直接的な指示方法と比較して、次の3つの利点があると考えられる。

モバイル端末の位置姿勢推定の誤差の影響を受けにくい

AR 技術を用いた直接重畳方式では、カメラ（視点）の位置姿勢推定が失敗し、アノテーションの位置合わせに誤差が生じると、アノテーションが指し示す位置がずれてしまい、誤った指示共有をしてしまう恐れがある。特に、屋外環境においてはモバイル端末の位置姿勢推定が難しいため、この問題が生じる可能性が高くなる。提案方式では、指示位置がジオラマモデルの座標系で表現されるため、常に正確な位置を提示し続けることができる。

ユーザの視野外の場所でも指示共有が可能

ユーザはジオラマモデルを眺めることで、ユーザの立つ場所からは見通せない位置を仮想的に視ることができる。タッチスクリーンを用いて、このジオラマモデル上の点を指示し、そこに矢印などのアイコンを表示することにより、指示位置や、互いが直接見通せない場所にいる（遠隔地にいる）ユーザ同士でも指示の共有が可能になる。

指示位置の奥行きを正確に認識できる

直接重畳方式では、異なる向きから指示位置を観察したい場合、ユーザが実空間中で移動して視点を変更する必要がある。提案方式にはこのような制限はなく、指示位置をまわりこむようにして見回すことで、指示位置の奥行きを正確に認識できる。

仮想環境における視点を操作する方法としては、一般的には、キーボードやマウス、ジョイスティックの入力によって行うことが挙げられる。これは、指先だけで仮想視点を自由な位置へ移動させることができるが、ユーザの向いている位置や方向と無関係な仮想視点にもできるため、ユーザは仮想視点の方向とユーザ自身が向いている方向の違いを理解する必要がある。地図とコンパスを見ながら目的地の方向を理解する時と同様に、人によってはこれが困難な場合がある。また、モバイル端末を持ちながらキーボードやマウスを操作すると両手を使う必要が生じる。

そこで、本論文では、仮想視点の向きをモバイル端末の向いている方向に合わせることでジオラマモデルと現実空間の方位を常に一致（整列）させて表示する手法を検討する。例えば、ユーザが端末を北に向けるとジオラマモデルも北を向く。ユーザがモバイル端末を動かすことで視点を操作できるため、片手で操作することも可能である。これにより目的地の方向判断がしやすく

なることは、空間認知の研究分野では整列効果として知られている。

3.2 指示位置の入力

提案インタフェースでは、指示位置はジオラマモデルの表面上にあることを前提とし、タッチスクリーンによる指示位置の入力を実現する。ユーザがジオラマモデルにタッチすると、レイキャストによって奥行きが設定される。

ユーザがモバイル端末のタッチスクリーン上の1点を入力すると、仮想視点を始点としてその点を通るレイとジオラマモデルとの交点の3次元位置が指示位置として設定される。

具体的な手順は、図 3-1 に示すように、ジオラマモデルを観察する仮想視点が A の位置・姿勢であるとき、仮想視点とタッチスクリーン上の位置を通る直線がジオラマモデルと交わる最初の点を指示位置とする。建物の陰など、現在の仮想視点 A からは見通すことが出来ない場所を指示する際は、指示者が仮想視点 B のように見通しの効く位置にモバイル端末をまず移動させてから指示位置を入力する。

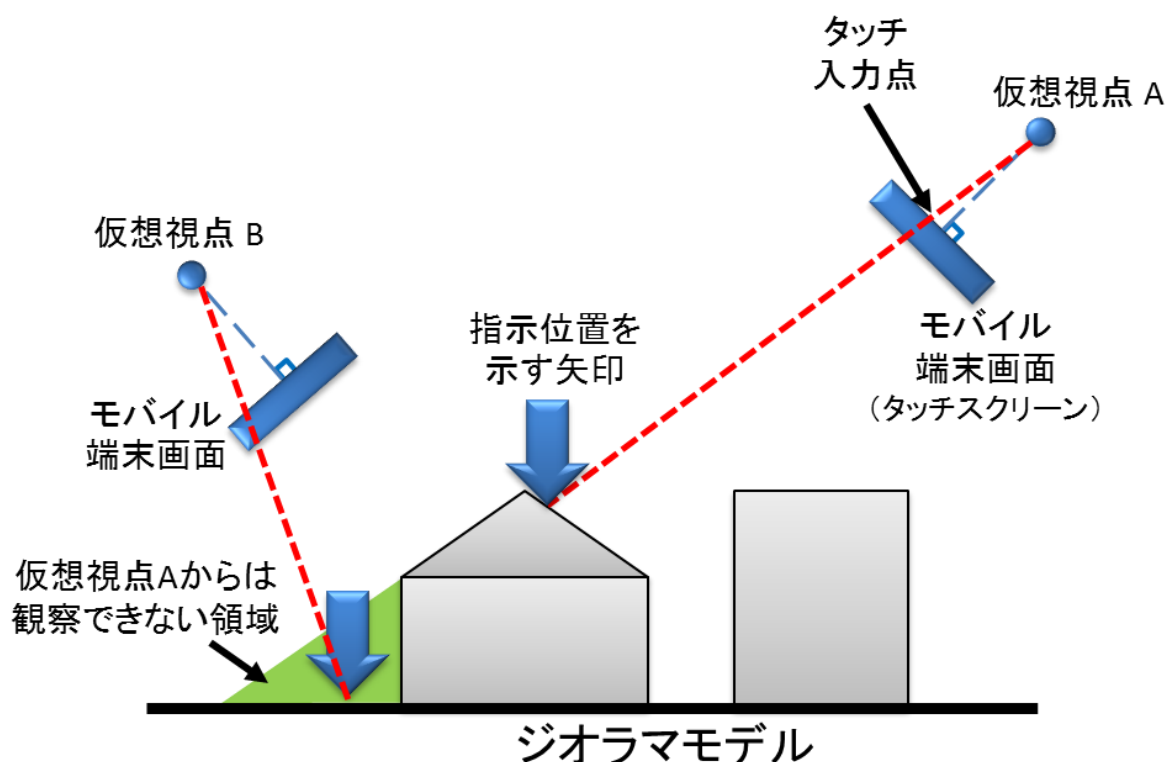


図 3-1 タッチスクリーンからのレイキャストを用いた指示位置の入力

3.3 指示位置の提示

提案インタフェースでは、図 3-2 に示すように、指示位置をジオラマモデル上に表示した矢印のアイコンで示す。それぞれのユーザからの指示位置を区別するために、矢印はユーザ毎に固有の色で表示する。

Virtual-Reality (VR)の研究において、3次元の指示位置 (Point of Interest) を提示するためのアイデアはいくつか提案されている [26] [35]。いずれの研究においても、指示位置を提示する際には、ユーザから指示位置が見えない場合への対処方法が重要となっている。提案インタフェースでは、ユーザによる仮想視点の変更と、遮蔽物の透過表示によってこの問題に対処する。

提案インタフェースでは、ユーザがジオラマモデルを観察する仮想視点进行操作することで、指示位置がどこにあっても、自由に仮想的に見ることができる。そのため、指示位置を自由に観察できる反面、指示位置が視野外にある場合には、ユーザはジオラマモデル上を探索し、指示位置が観察できる仮想視点を見つけ出す必要がある。本論文における提案インタフェースの実装では、タッチスクリーン操作によってジオラマモデルの拡大／縮小を可能にし、指示位置が見つからなければジオラマモデルを縮小することで広域を観察できるようにし、指示位置を見つけやすくする。

仮想視点から観察できる範囲に指示位置があったとしても、状況によっては矢印の示す先がわからにくくなる場合がある。例えば、ジオラマモデル上の建物によって矢印が隠されて見えない場合、ユーザは仮想視点を動かして、矢印の位置を探索する必要がある。それを容易にするために提案インタフェースでは、指示位置とジオラマモデルの建物が重なってしまう場合に、ジオラマモデルの建物を透過させて表示する (図 3-3a)。これは、Elmqvist ら [36]による、遮蔽された物体の表現方法の分類では、遮蔽物を透かす“Virtual X-ray”に相当する。この方法は、仮想視点の制御をユーザが行える上に、複数の視点や、幾何的な歪みが必要ないため、指示位置の存在を簡単に理解できる方法であると考えられる。

図 3-3 (b)のように、矢印が示す先にジオラマモデルの建物が一つしかない場合、指示位置はその建物上の点を示している。一方、矢印が示す先にジオラマモデルの建物が複数見える場合、固定視点では指示位置がどちらの建物上の点を示しているのか判断できない (図 3-3c)。この場合でも、ユーザは仮想視点を操作して異なる視点から視ることで、指示位置がどの建物上の点を示しているのかを確認することができる。図 3-3 (c)の例では、図 3-3 (c')のように少し仮想視点を横に動かすことで、視差から矢印の奥行きを知覚することができ手前の建物の角を指示されていることがわかる。



図 3-2 指示共有の実現

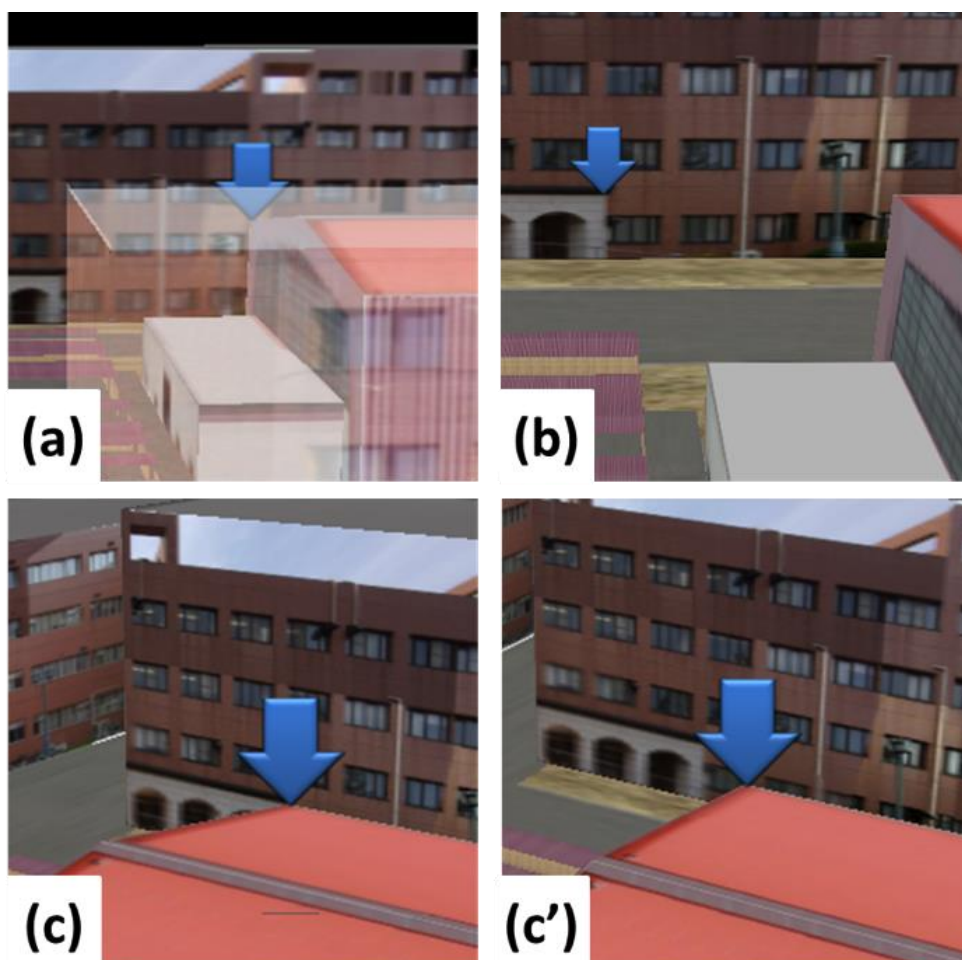


図 3-3 ジオラマモデル上の矢印による指示位置の提示例

3.4 実写映像への重畳表示

本節では、一般的なビデオシースルー型の AR/MR システム [3]で採用されている、モバイル端末のカメラで撮影した実写映像にジオラマモデルを重畳表示する方式（図 3-4 左）の是非を検証する。

複合現実型ジオラマインタフェースでは、ユーザがモバイル端末のカメラと同じ方向を向いているときに、ユーザの見ている実際の風景とモバイル端末の画面に提示されるジオラマモデルとを見比べることができる。

原則としては、ジオラマモデルを実写映像へ重畳表示することにより、画面内で屋外環境とジオラマモデルを見比べることができるため、画面から目を離す必要が減り、ジオラマモデル上の矢印が示す指示位置がわかりやすくなると考えられる。しかし、モバイル端末は画面のサイズが小さいことや、屋外での使用時には画面のコントラストが低くなる場合があることから、実写映像とジオラマモデルの境界が見分けにくくなる場合があり、必ずしも指示位置が見やすくなるとは限らない。

画面にジオラマモデルのみを表示し、実写映像へは重畳しない場合、重畳表示をする場合に比べて、画面に提示される情報が限定されるため、ジオラマモデルの観察に集中しやすくなると考えられる。しかし、実際の風景とジオラマモデルとを見比べるためには、画面から目を離して実際の風景を見なければならないという手間がある。

そこで、被験者実験を通じて両方式の特性を比較することにより、ユーザにとって理解しやすい表示方式を検討する。

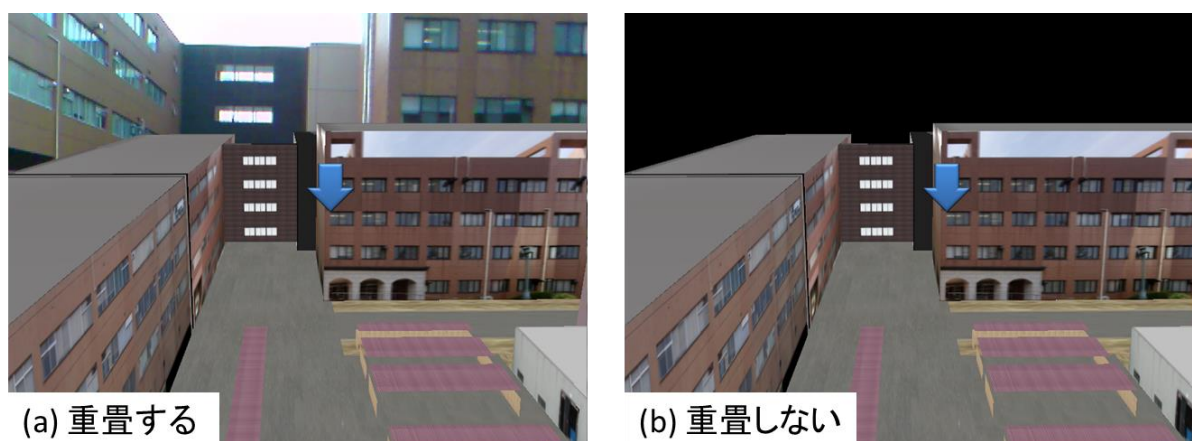


図 3-4 実写映像への重畳表示の有無

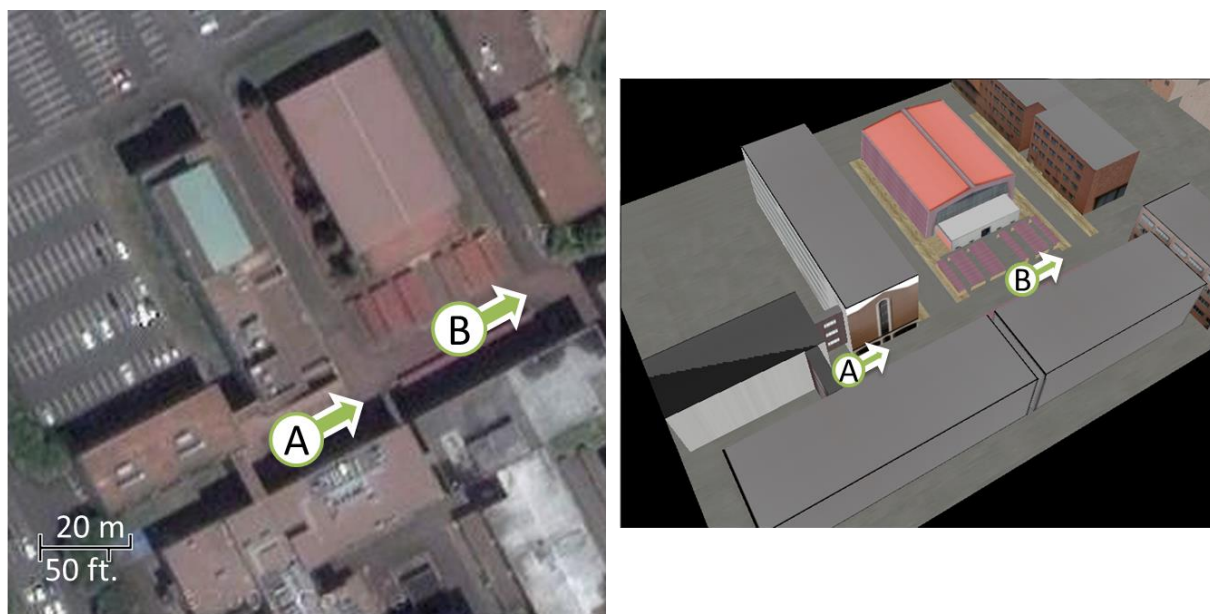
3.4.1 実験環境

ジオラマモデルを実写映像へ重畳する方式の是非を検証するため、屋外環境において被験者実験を行った。

本実験で用いる複合現実型ジオラマインタフェースは、4.5型のタッチパネルディスプレイを搭載した SONY 社の小型ノート PC (VGN-UX92NS) と、InterSense 社の 3 軸慣性センサ InertiaCube3 を用いて実装した。複合現実型ジオラマインタフェースの画面構成は、実写映像へ重畳する場合は、4.2 節の図 4-2 に示すものを用いた。実写映像へ重畳しない場合は、メインウィンドウ部の表示をジオラマモデルのみとし、それ以外の画面構成は同一とする。

実験環境として、図 3-5 (左) に示す筑波大学構内の一部 (120[m]x100[m]程のエリア) を利用し、図 3-5 (右) に示すジオラマモデルを作成した。被験者の立ち位置を図中の A 地点とし、図 3-6 に示す 20 箇所を指示位置として利用した。これら指示位置は、被験者が移動せずとも見えるように設定した。

A 地点は三方が建物で囲われた建物の密集する地点で、ジオラマモデル上で矢印が建物に遮蔽されやすいため、建物ごとに透過して表示する図 3-3 (a) のような状況が発生しやすい環境である。指示位置が分かりにくい状況として、図 3-3 (c) のように、矢印が示す先にジオラマモデルの建物が複数見える状況も実験に含めた。建物の窓枠のように見た目の類似した場所が周囲に複数ある箇所も指示位置として指定した。



©2008 Google ©2008 DigitalGlobe, GeoEye

図 3-5 実験環境の航空写真 (左) と実験に用いたジオラマモデル (右)

◎先に行う10タスク △後に行う10タスク (番号は提示順序)

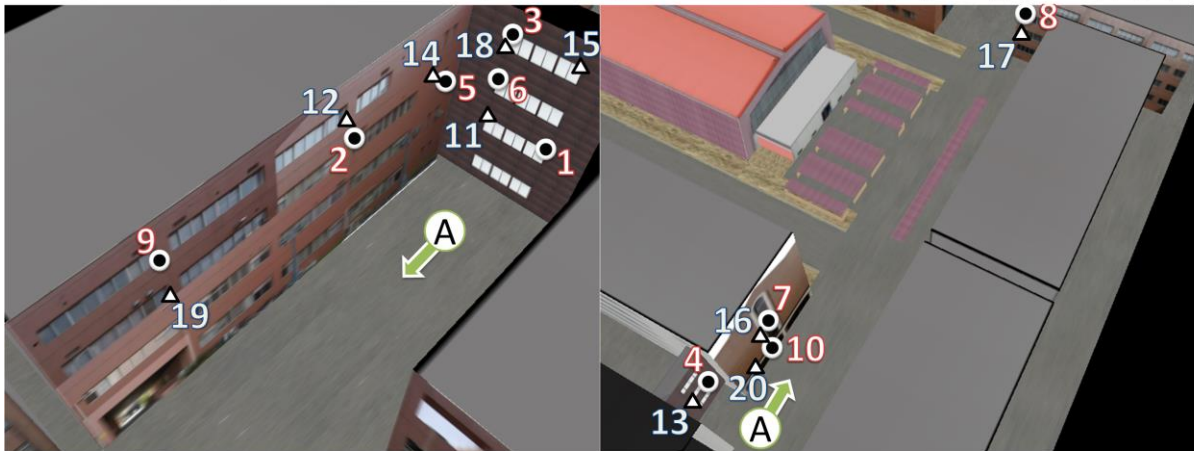


図 3-6 A 地点で行う写真撮影タスクに用いる指示位置

3.4.2 ユーザタスク

本実験で被験者に課すタスクには、ジオラマモデル上に表示された矢印の示している位置がどこであるかを理解し、実際のその場所にカメラを向けて写真を撮るという、写真撮影タスクを用いる。

このタスクでは、被験者は図 3-5 中の A 地点に立ち、準備ができればモバイル端末の画面にタッチする。次に複合現実型ジオラマインタフェースが画面に提示され、ジオラマモデル上に青い矢印で指示位置が提示される。被験者はその指示位置が現実空間中のどこであるかを理解し、そこに向けて写真を撮影する。写真撮影に用いるカメラはモバイル端末に組み込まれたものを使い、モバイル端末のスペースキーを押して撮影する。

タスクの所要時間を計測しているため、被験者にはできるだけ早くタスクを終えるよう指示する。また、移動はせずその場で向きだけ変えて撮影をすること、タスクを始める前に必ず決められた方向（図中に矢印で示す方向）に向くことを指示する。

3.4.3 実験手順

実験では各被験者に対し、ジオラマモデルを実写映像へ重畳するインタフェースと、重畳しないインタフェースの 2 種類を提示する。被験者は、それぞれのインタフェースで操作に慣れるまで数回練習し、写真撮影タスクを 10 回繰り返す。その後、主観的な印象を問う質問を行う。

インタフェースの提示順序の影響を考慮し、被験者を、先に重畳する方式のインタフェースを用いるグループと、先に重畳しない方式のインタフェースを用いるグループの 2 つに分けて実施する。

主観的な印象を問う質問は、表 3-1 に示す 4 問とし、どちらの方式が優れていると思うかを尋

ねた。評価はいずれも 5 段階のリッカート尺度である。ジオラマモデルを実写映像へ重畳する場合を方式 A、重畳しない場合を方式 B とし、明らかに方式 A だと思う場合は 1、どちらかと言え
ば方式 A ならば 2、わからない、どちらでもない場合を 3、どちらかと言え方式 B ならば 4、
明らかに方式 B ならば 5 として回答させた。

表 3-1 主観的な印象に関する質問

番号	質問
Q3-1	指示された場所がわかりやすいのはどちらでしたか？
Q3-2	どちらの方が思い通りの視点にしやすいかったですか？
Q3-3	今後使うとしたら、どちらを使いたいですか？
Q3-4	どちらの方が見やすい映像だと思いますか？

3.4.4 実験結果

実験環境をよく知る 20 歳代の男子学生 9 名を被験者として被験者実験を行った結果を示す。

図 3-7 は各質問に対して、1～5 の評点を付けた人数の割合である。実写映像上へジオラマモデルを重畳表示する方式が印象評価に与える影響は、いずれの設問に対しても被験者らの意見は分かれていた。Wilcoxon の符号付き順位和検定の結果、各質問の有意確率 p は、Q3-1: $p=0.847$, Q3-2: $p=0.992$, Q3-3: $p=0.502$, Q3-4: $p=0.441$ となり、いずれも有意差は見られなかった。

表 3-2 に示すように、タスク所要時間の平均にも有意差は見られなかった。図 3-8 に所要時間の箱ひげ図を示す。

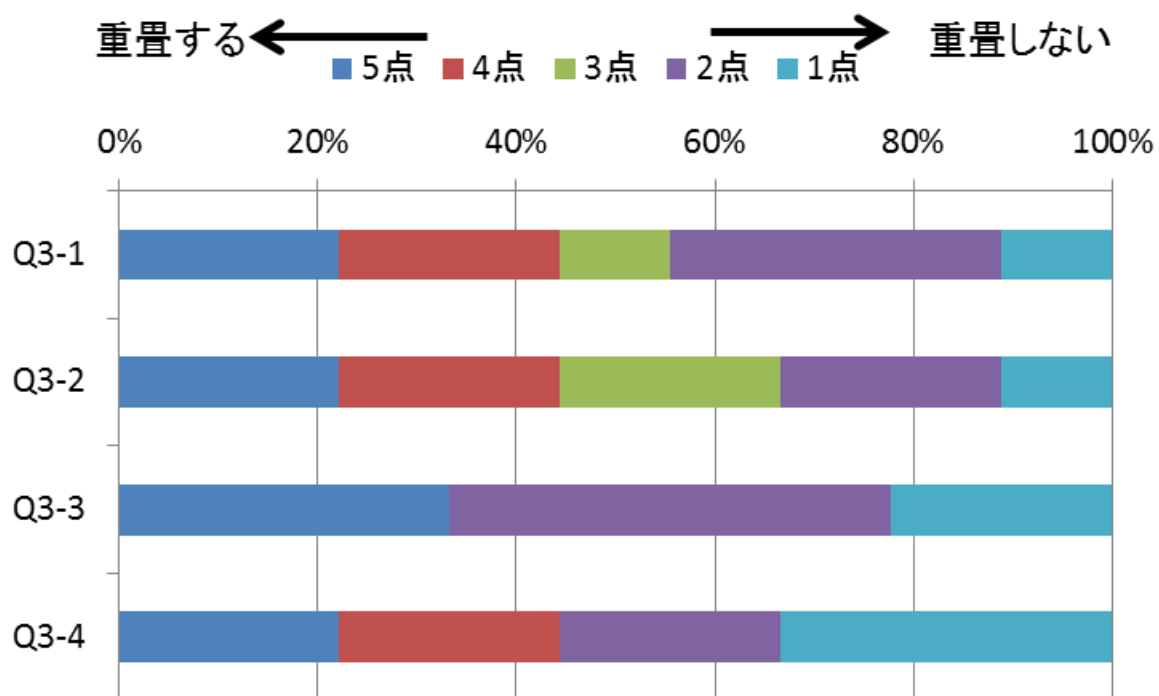


図 3-7 実写映像への重畳表示の有無の比較実験 印象評価結果

表 3-2 実写映像への重畳表示の有無の比較 タスク所要時間の平均

実験場所	重畳する	重畳しない	p 値(対応のある t 検定)
A 地点	9.9 秒(4.5 秒)	10.8 秒(4.2 秒)	p=0.357

()内は標準偏差.

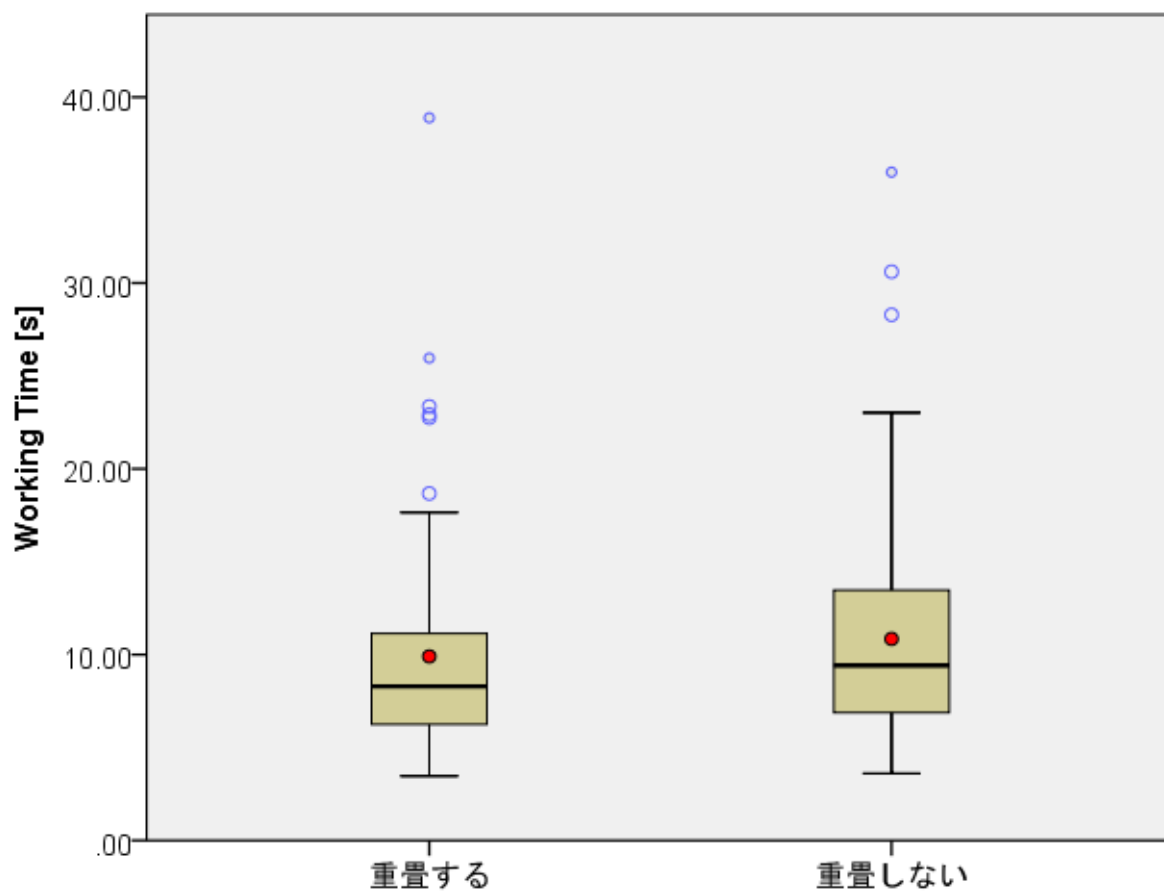


図 3-8 実写映像への重畳表示の有無の比較 タスク所要時間
赤点は平均値を、青丸は外れ値を示す。

表 3-3 実写映像への重畳表示の有無の比較 タッチパネル操作の利用状況

利用状況	拡大／縮小	視点の水平移動
使用せず	3 名	2 名
重畳する	0 名	2 名 (1.0)
重畳しない	1 名 (2.0)	2 名 (1.0)
両方で使用	5 名 (あり:5.6, なし:3.6)	3 名 (あり:3.0, なし:3.0)

()内は 10 タスク中の平均利用タスク数

3.4.5 考察

図 3-7 に示すように、実写映像上へジオラマモデルを重畳する方式が指示位置を示す矢印の見やすさや、提案インタフェースの使いやすさに貢献していると考えている被験者は、全体の半数程度であった。

重畳表示する方式に対する好意的なコメントとしては、「縮小表示だと大まかな方向しかわからないため、知らない場所では苦労しそうだと感じた」など、GUI 部の縮小表示よりも大きく実写映像を観察できる点を評価する意見がほとんどであった。

表 3-2 や図 3-8 に示すように、1 回のタスクに要する所要時間の平均は、重畳する方式の方が短くなる傾向が見られたものの、有意な違いは見られなかった。また、タッチパネル操作に関しても表 3-3 に示すように、両方式で大きな違いは見られなかった。以上より、実写映像上へジオラマモデルを重畳表示する方式は、指示位置の見やすさや提案インタフェースの使いやすさに貢献していると考えているのは半数程度であった。

実写映像の提示に否定的な被験者らのコメントとしては、「画面中でジオラマモデルが占める面積が大きい時に実写があまり見えない」「ジオラマモデルと実写映像の境界がわかりづらい」など、提示方法に関する具体的な問題点の指摘があった。Q3-3「今後使うとしたら、どちらを使いたいですか？」という質問では、重畳しない場合を選んだものの、「重畳する場合はジオラマモデルと屋外環境を見比べやすかった」とコメントした被験者もいた。このことから我々は、実写映像とジオラマモデルの境界線をコントラストの高い色で強調表示してわかりやすくすることや、実写映像とジオラマモデルが画面中に占める面積比などに注意することで、重畳する場合を指示位置の見やすさ向上に結び付けられるよう方式を改善できる可能性があると考えている。

第4章 複合現実型ジオラマインタフェース による指示共有

本論文で提案する複合現実型ジオラマインタフェースは、3次元の指示位置の入力・提示を行うインタフェースである。このインタフェースを用いて、指示共有を行うためには、位置情報の記録を行うサーバとの、無線ネットワークを介した通信や、指示位置以外の情報を提示するための画面表示が必要である。本章では、それらの詳細について述べ、提案インタフェースを用いた指示共有の実現方法の一例を示す。

4.1 遠隔指示共有

指示共有を行うための前提として、指示共有を行う人々（ユーザ）がそれぞれモバイル端末を持っており、その画面上には複合現実型ジオラマインタフェースが提示されているものとする。全てのユーザに提示されているジオラマモデルは共通のものであるが、ジオラマモデルを観察する仮想視点は、それぞれのモバイル端末の位置姿勢によって異なる。

図 4-1 に示すように、あるユーザから他のユーザへの指示は、以下の手順で行う。

- ① あるユーザがジオラマモデルにタッチすると、3.2 節に示す方法で入力された3次元の指示位置が、無線ネットワークを介してサーバに送信される。
- ② サーバは、ユーザからの指示位置を受信すると、他のユーザにその指示位置の情報を送信する。
- ③ 他のユーザのジオラマモデル上にその指示位置が矢印で表示されることにより、指示共有が実現できる。

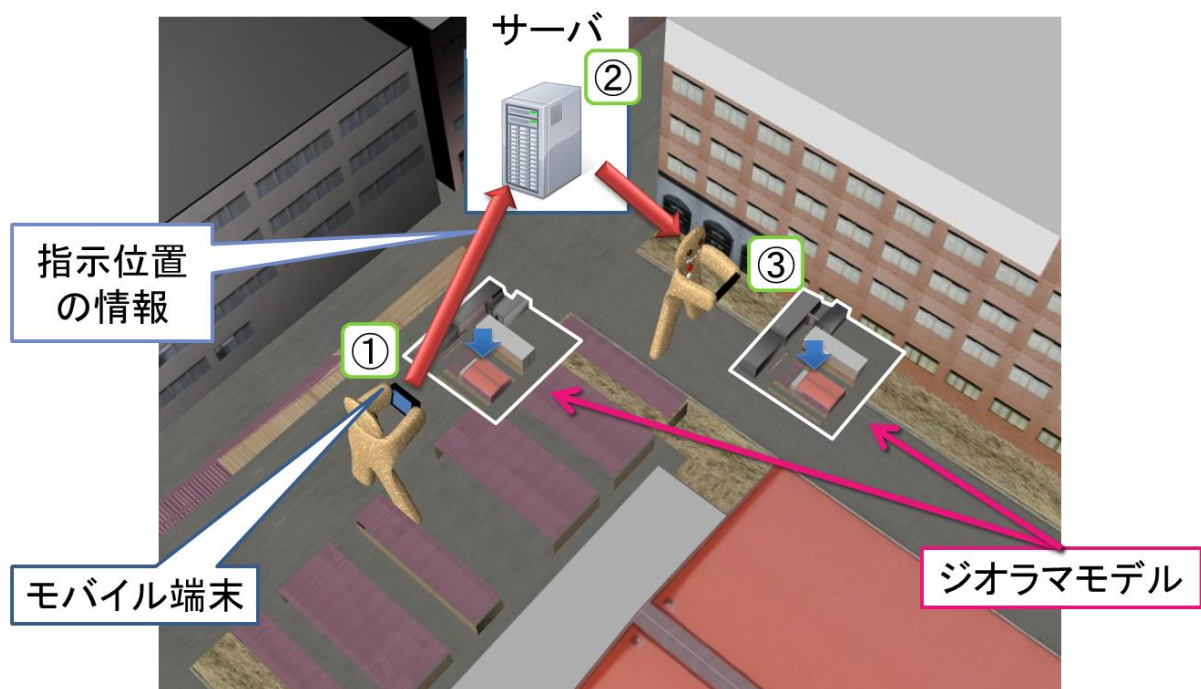


図 4-1 無線ネットワークを介した提案インタフェースによる指示共有方式

4.2 画面の構成

指示共有を行う際には、位置情報とその他の情報を提示する必要がある。本論文では、複合現実型ジオラマインタフェースを提示するメインウィンドウと、その他の情報を提示する GUI（グラフィカルユーザインタフェース）部に分割した画面構成を用いる。

図 4-2 に、シングルタッチ（入力点が 1 点のみ）のタッチスクリーンを搭載したモバイル端末向けの画面構成例を示す。

ジオラマモデルの拡大／縮小を制御するスライダー、所要時間、カメラで撮影した実写画像を縮小表示する GUI 部からなる。ジオラマモデル上には、指示位置を示す矢印と、ユーザの現在地がアイコンで表示される。ユーザはスライダーをタッチスクリーンで操作し、ジオラマモデルを拡大／縮小して視ることが出来る。

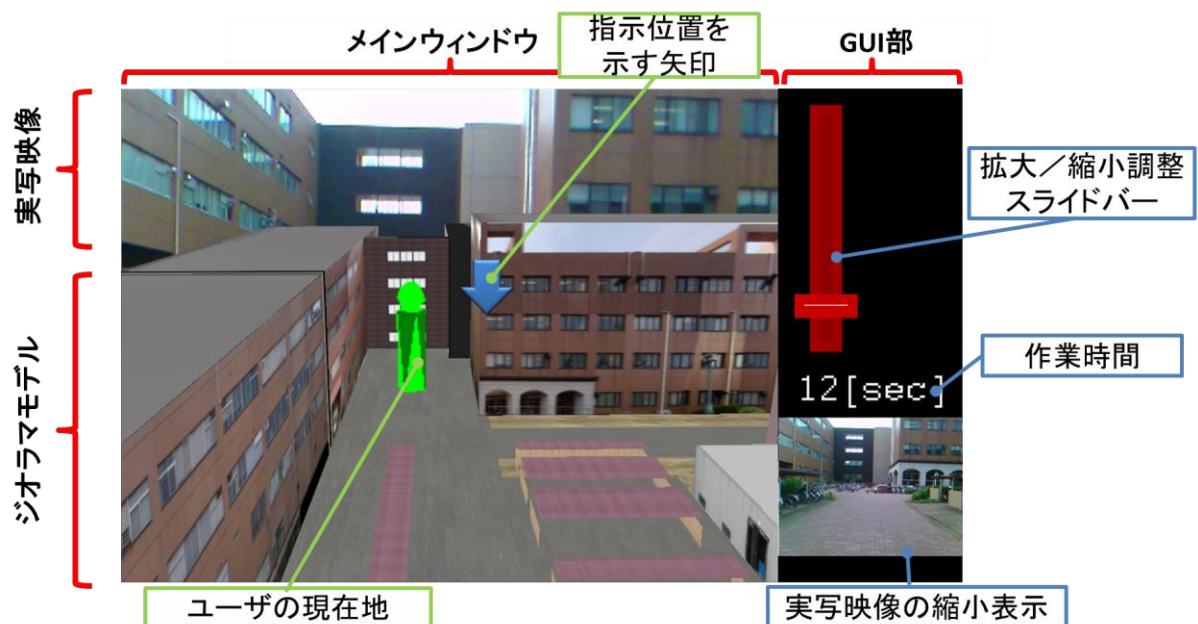


図 4-2 指示共有システムの画面構成

4.3 音声指示の併用

一般的に位置情報の指示共有を行う際には、「ここに注目してください」「ここに置いて下さい」というように、その位置を指示した意図の説明や、「2階の左から3番目の窓」というように、指示位置を特定するための詳細な説明を付加するが多い。

本論文で提案する複合現実型ジオラマインタフェースは、ジオラマモデルを用いて3次元の指示位置を明確に入力・提示することができるため、指示共有を行う際に、指示位置を特定するための言葉による説明を省略することができる。また、意図の説明については、ジオラマモデル上に提示するアイコンの種類を複数用意し、それぞれの意味を共同作業者と決めておくことで、予め決められた種類の意図を伝えることができる。これらのことから、提案インタフェースを用いることで、指示共有の際に音声指示の必要性を軽減することが可能であると考えられる。

しかしながら、共同作業者の感情などを伝えるなど、指示共有を円滑に行うためには音声指示は依然として必要不可欠であると考えられるため、提案インタフェースを実際に共同作業で用いる際には、音声指示との併用を前提として考える。

本論文で行う被験者実験では、複合現実型ジオラマインタフェースの評価を行うために、音声指示による指示位置の説明や、意図の説明は行わない。

第5章 ジオラマモデルの視点操作

3次元のジオラマモデルを導入した場合、いかなる仮想視点を用いて2次元の画面上へ投影するかが問題となる。本論文では、モバイル端末の姿勢を用いることにより、ユーザが簡単に仮想視点を操作できることと、現実空間における指示位置がわかりやすいことの両立を目指す。

5.1 整列効果

人が地図を読む際の一般的な性質として、整列効果というものが知られている。これは、地図と実際の空間との方位が一致している場合に目的地の方向判断が容易になるという効果で、3次元地図では整列効果の影響がより強くなる [37]。一般的なカーナビゲーションシステムでは、進行方向に合わせて地図を回転させる方式（整列方式）だけでなく、地図の上方向を常に北向きに固定する方式（非整列方式）を選択して利用することもできる。

本節では、整列効果を利用するために、仮想視点の向きをモバイル端末の向いている方向に合わせることで、図 5-1 に示すようにジオラマモデルと現実空間の方位を常に一致させて表示する整列方式の是非を検証するために、非整列方式との比較実験を行う。

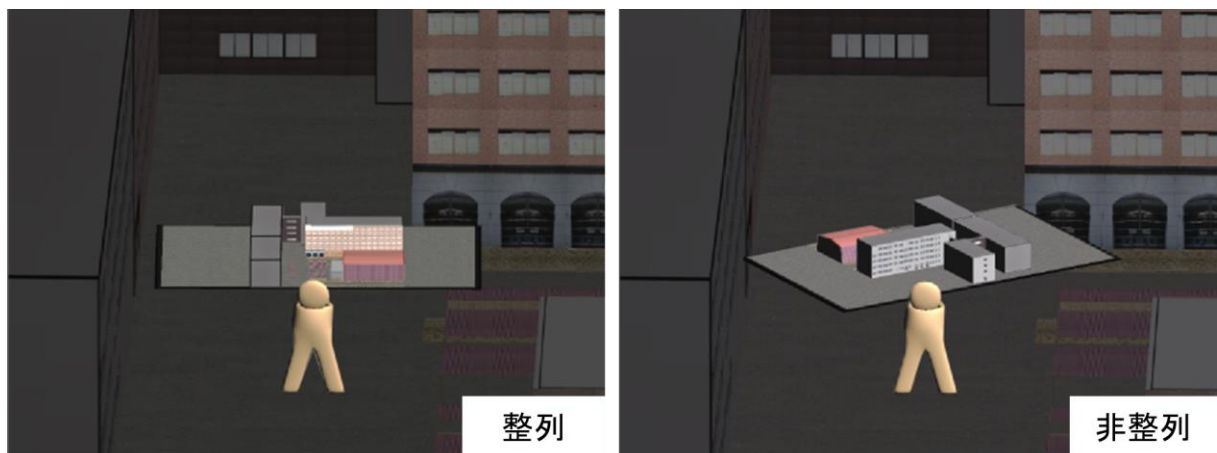


図 5-1 整列方式と非整列方式

5.1.1 整列／非整列方式の比較

整列方式の場合、ジオラマモデルと現実空間の方位を合わせるのに必要な認知過程である、心的回転（メンタルローテーション）に要する時間が削減され [38]、方向判断が素早くなると考えられる。しかし、仮想視点の向きによっては、指示位置を示す矢印が視野外になる場合や、図 3-3 (c)のように手前と奥の建物のどちらを指示されているのかが不明瞭になる場合があり、指示位置を正しく把握できる仮想視点をユーザが操作して見つけ出す必要がある。

一方、非整列方式の場合、モバイル端末がどのような姿勢にあっても一定の見え方でジオラマモデルを利用できるという利点がある。しかし、ユーザがジオラマモデルと現実空間の方位の違いを理解する必要がある。

以上のように、ジオラマモデルを整列／非整列させる方式にはそれぞれ利点と欠点があるため、被験者実験による比較を行う。

3.4 節の実験と同じく、モバイル端末として SONY 社の小型ノート PC (VGN-UX90PS)、3 軸慣性センサとして InterSense 社の InertiaCube3 を用いた。実験に使用した機器の使用は付録に示す。実験環境は、図 5-2 に示す筑波大学構内の C 地点である。

本実験で被験者に課すタスクは、3.4.2 節で示した写真撮影タスクである。

被験者は C 地点に立ち、ジオラマモデル上に表示された矢印の示している位置がどこであるかを理解し、実際のその場所にカメラを向けて写真を撮る。このタスクを整列方式で 5 回、非整列方式で 5 回実行した後、表 5-1 に示す質問に答える。

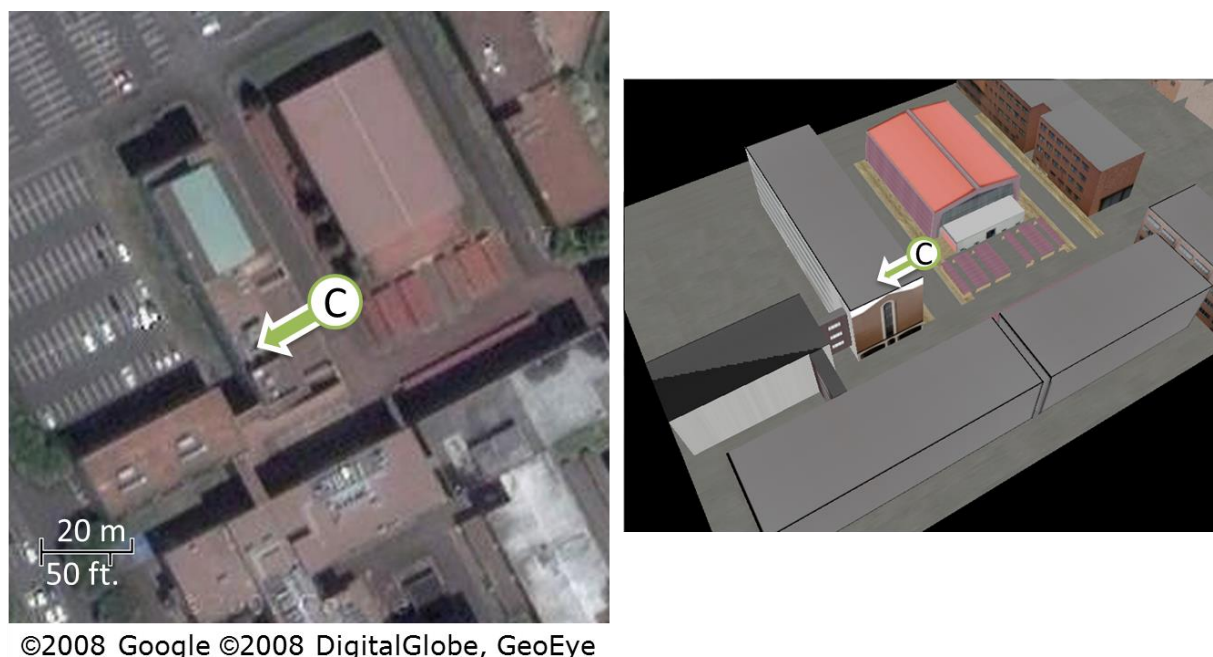


図 5-2 実験環境（C 地点）の航空写真（左）と実験に用いたジオラマモデル（右）

表 5-1 整列方式／非整列方式の比較 主観的な印象に関する質問

番号	質問	尺度
Q5-1	どちらの方が見やすかったですか？	整列方式 (2) ～ 非整列方式 (-2)
Q5-2	どちらの方が指示位置が わかりやすかったですか？	整列方式 (2) ～ 非整列方式 (-2)

5.1.2 実験結果と考察

実験環境をよく知る 20 歳代の男子学生 7 名を被験者として実験を行った。その結果、表 5-1 に示す質問 Q5-1 の平均点は 1.4 (分散:0.8)、Q5-2 の平均点は 1.1 (分散:0.9) となり、ジオラマモデルを整列させる方式のほうが、見やすく、指示位置がわかりやすいという結果が得られた。

タスク所要時間の平均を表 5-2 に、箱ひげ図を図 5-3 に示す。非整列方式のほうが短い傾向があるが、有意差はなかった。この理由として、非整列方式では、被験者による仮想視点の操作が必要ないことが考えられる。しかし、被験者による仮想視点の操作が必要な整列方式においても有意な時間の増加は見られなかった。

表 5-2 写真撮影タスク所要時間の平均

実験場所	整列	非整列	p 値(対応のある t 検定)
C 地点	11 秒(6.5 秒)	10.2 秒(5.0 秒)	p=0.24

()内は標準偏差。

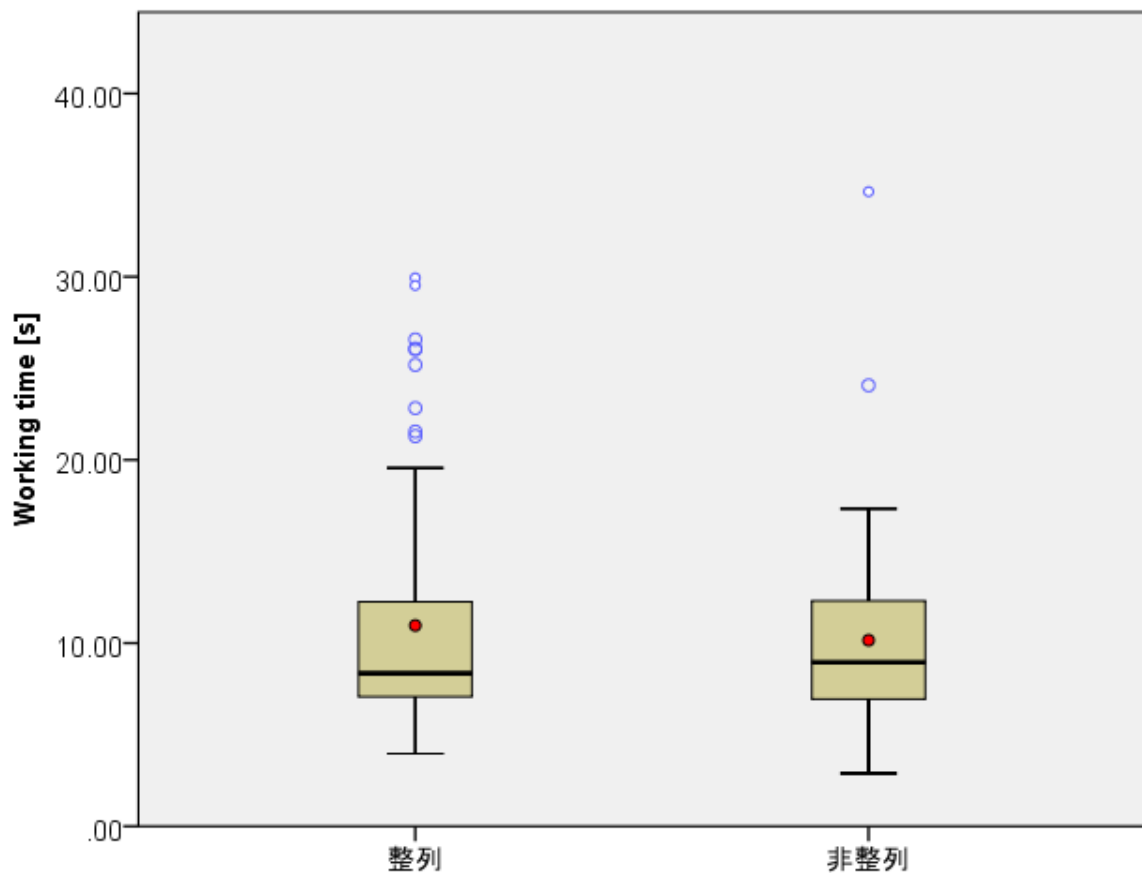


図 5-3 整列方式／非整列方式の比較 タスク所要時間

5.2 整列性を維持した視点操作方式

本節では，ジオラマモデルの視点操作方式として，実物のジオラマモデルと同様の見え方を再現する Augmented-Reality 型，視点の平行移動をタッチパネルのドラッグによる手動操作のみで行う Pan/tilt 方式，タッチパネルのドラッグ操作による視点の平行移動に加えてジオラマモデル上の注視点が常に視野の中心に配置される Dolly-round/crane (D/C)方式とを取り上げる．

5.2.1 Augmented-Reality (AR)方式

AR 方式は，ジオラマモデルの位置・姿勢を現実空間中に固定し，仮想視点をモバイル端末に固定することで，あたかもテーブルの上に置かれたジオラマモデルを，モバイル端末画面をのぞき窓として観察しているような見え方を再現する視点操作方式となる（図 5-4）．モバイル端末の

位置・姿勢を用いて 6 自由度の視点操作を行える。AR 方式は、実物の建築模型と同様の見え方をするという実世界のメタファを持つため、拡張現実感を用いたオーソドックスな仮想視点の操作方式と考えられ、本章で比較する 3 方式の中では最も直観的な操作方法であると予想される。

しかし、AR 方式の場合、図 3-3 (c)に示した例のように、ある視点からの映像ではジオラマモデル上の矢印の示す位置が曖昧な状況においては、その矢印を異なる視点からも視ることで、指示位置がどの建物上の点を示しているのかを確認する必要がある。そのためには図 5-5 (a)に示すように、ユーザがモバイル端末を回転・平行移動させる必要があり、ユーザがジオラマモデルを回りこむように移動する必要がある。

本論文では、AR 方式の比較対象として、モバイル端末の平行移動を入力として使用せず、図 5-5 (b)のように、現実空間におけるジオラマモデルの提示位置を平行移動する方式として、Pan/tilt 方式と、Dolly-round/crane (D/C) 方式を挙げる。これらの方式では、ユーザが移動しなくとも、ジオラマモデルを回りこむような仮想視点の操作が行える。

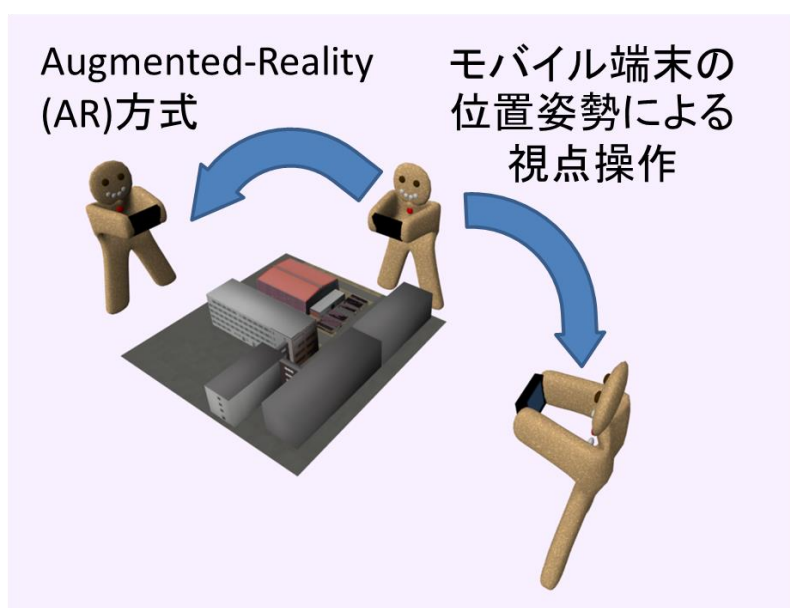


図 5-4 Augmented-Reality (AR)方式

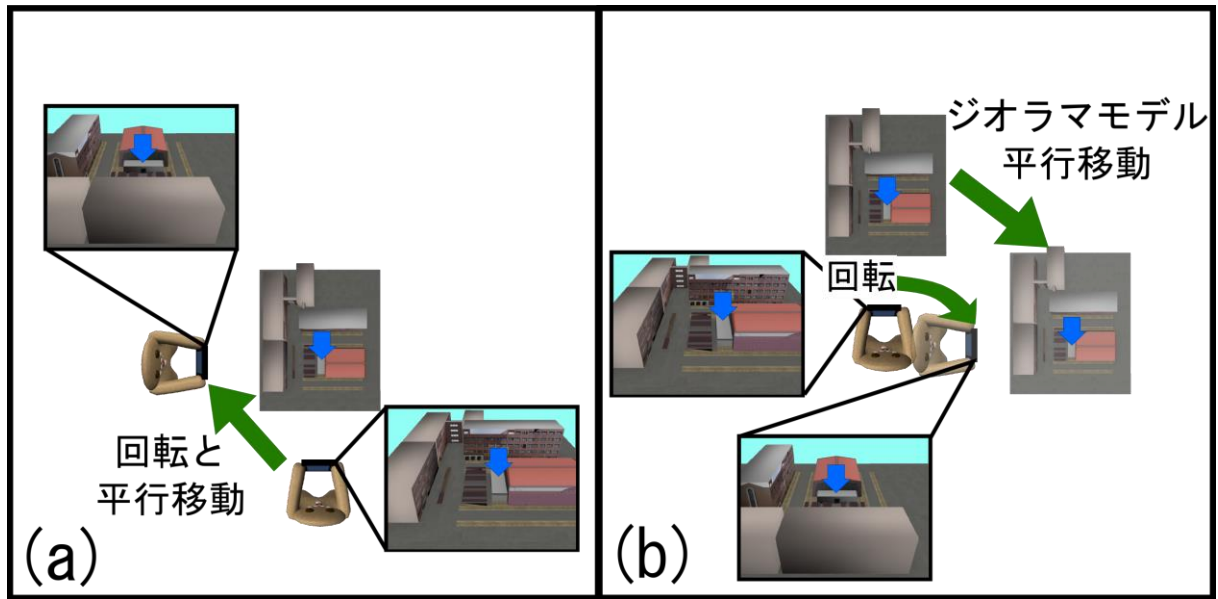


図 5-5 視点の回りこみ操作例

5.2.2 Pan/tilt 方式

この方式は、モバイル端末の位置を仮想視点の位置とする代わりに、タッチパネルのドラッグ操作を用いて仮想視点の水平位置を操作するという点を除いては、AR 方式と同じ操作方式である。AR 方式と比較して、ユーザは視点の平行移動のために実際にモバイル端末を移動させる必要がないという利点がある。

図 5-6 に示すように、ユーザがモバイル端末を動かしても、仮想視点の位置は変わらないため、タッチスクリーンをなぞることによってジオラマモデルを観察しやすい位置（正面）に仮想視点を水平に平行移動させる。

仮想視点は、タッチスクリーンを左から右へなぞることで左方向に水平移動し、上から下へなぞることで前方へ水平移動する。この方式では、垂直方向への仮想視点の移動は行えないため、5 自由度の視点操作となる。

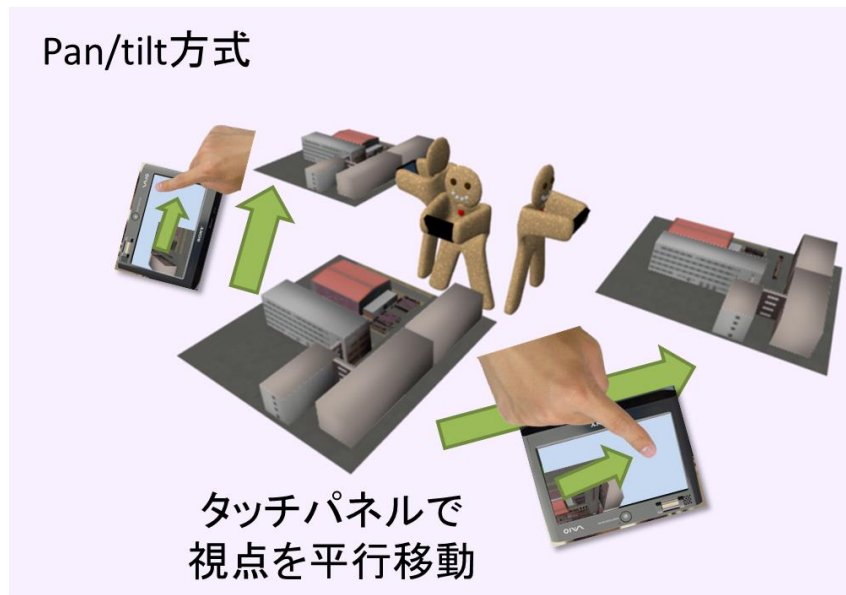


図 5-6 Pan/tilt 方式

5.2.3 Dolly-round/crane (D/C)方式

D/C 方式は、Pan/tilt 方式の操作方式に加えて、モバイル端末の姿勢に応じた自動的な仮想視点の水平移動を行う。これは、モバイル端末の姿勢に応じて仮想視点ジオラマモデルを中心とした仮想的な球面上を移動する方式である。画面の中央には常に球面の中心が表示されており、タッチスクリーンのドラッグ操作によって球面の中心位置をジオラマモデル上で水平移動させることができる。球面の中心を指示位置の付近に設定すれば、モバイル端末の向きを変えることにより様々な仮想視点で指示位置を観察できる。

この視点操作方式は、Orbital viewing と呼ばれており、仮想物体を観察するための視点操作方式としては操作しやすいものであると言われている [39]。Koller ら [40]は、Orbital viewing は、ジオラマモデルや3次元地図の提示に特に適しているとしているが、その理由として整列効果が期待できることを挙げている。本論文では、整列効果が期待できる3つの視点操作方式を比較し、視点操作のしやすさと指示位置のわかりやすさを調べる。

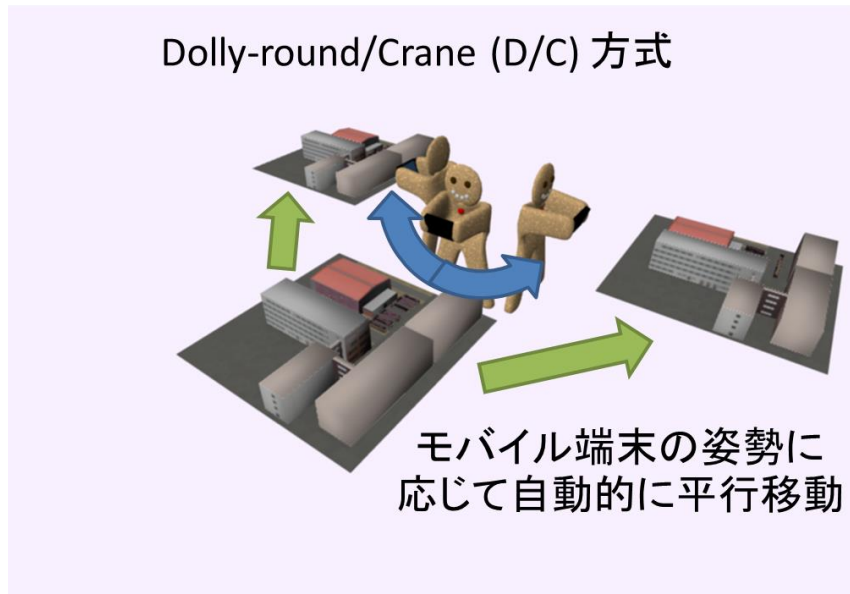


図 5-7 Dolly-round/crane (D/C)方式

5.3 整列性を維持した視点操作方式の比較実験

本節では、整列性を維持した仮想視点の操作方式として、5.2 節で示した Augmented-Reality (AR)方式、Pan/tilt 方式、D/C 方式をとりあげ、屋外環境における被験者実験によって視点操作のしやすさと指示位置のわかりやすさを比較する。実験では、被験者に写真撮影タスクを実行させ、使いやすさに関する印象評価と、タスクの所要時間の計測を行う。タスクの所要時間は、ジオラマモデル上の矢印が示す指示位置に対応する点を現実空間中で見つけるのに要する時間として考える。

5.3.1 実験環境

本実験で用いる複合現実型ジオラマインタフェースには、4.5 型のタッチパネルディスプレイを搭載した SONY 社の小型ノート PC (VGN-UX92NS) を用い、4.2 節に示す画面構成を用いた。AR 方式におけるモバイル端末の位置・姿勢推定には、三脚で地上から 75[cm]の高さに置いた 15[cm] x 15[cm]の AR マーカ [41]を用いる。Pan/tilt 方式および D/C 方式におけるモバイル端末の姿勢推定には、InterSense 社の 3 軸慣性センサ InertiaCube3 を用いた。

本実験では、3.4 節の実験と同じく、図 3-5 に示す筑波大学構内の一部 (120[m]x100[m]程のエリア) を屋外環境としてジオラマモデル化した。実験場所は、図 3-5 中に A 地点として示す位置とした。

5.3.2 ユーザタスク

写真撮影タスクを遂行するためには、被験者はジオラマモデル上に表示された矢印の示している位置がどこであるかを正しく理解し、その位置にカメラを向けて写真を撮影する必要がある。

被験者には、タスクの所要時間を計測しているため、できるだけ早くタスクを終えるよう指示する。また、移動はせずその場で向きだけ変えて撮影をすること、タスクを始める前に必ず決められた方向（図 3-5 の A 地点に矢印で示す方向）に向くことを指示する。

5.3.3 実験手順

実験では、被験者は仮想視点の操作方式が異なる 2 種類の複合現実型インタフェースを用いて写真撮影タスクを実行する。それぞれインタフェースで操作に慣れるまで数回練習し、写真撮影タスクを 10 回繰り返す。被験者はその後、質問紙による印象評価を行う。このような一対比較を、5.2 節で示した仮想視点操作方式の全ての組み合わせで行う。組み合わせの順序は被験者ごとに変更する。

5.3.4 実験結果

図 5-8 に主観的な印象評価の結果を示す。被験者は 20 歳代の学生 17 名（男性 15 名女性 2 名）である。質問は表 3-1 に示すうち Q3-1～Q3-3 の 3 問とした。評価はいずれも、5 段階のリッカート尺度である（例えば方式 A と B の比較の場合、1:明らかに方式 A、2:どちらかと言えば方式 A、3:わからない、どちらでもない、4:どちらかと言えば方式 B、5:明らかに方式 B の 5 段階）。

シェッフェの一対比較（中屋変法）による検定の結果、D/C 方式は AR 方式に比べ、全ての質問において有意に優れていた（有意水準 1%）。また、Pan/tilt 方式は質問 Q3-3 においてのみ、AR 方式より有意に優れていた（有意水準 5%）。その他の結果に有意差はみられなかった。

図 5-9 にタスクの所要時間の計測結果を示す。被験者は 20 歳代の学生 15 名（男性 13 名女性 2 名）である。1 次元配置分散分析の結果、各操作方式間の平均値に有意水準 5%の有意差があった ($F(2,925)=4.797, p=0.008$)。多重比較の結果、D/C 方式（平均 12.2[sec]）では、所要時間が AR 方式（平均 13.6[sec], $p=0.016$ ）や Pan/tilt 方式（平均 13.7[sec], $p=0.030$ ）に比べて有意に短かった（有意水準 5%）。

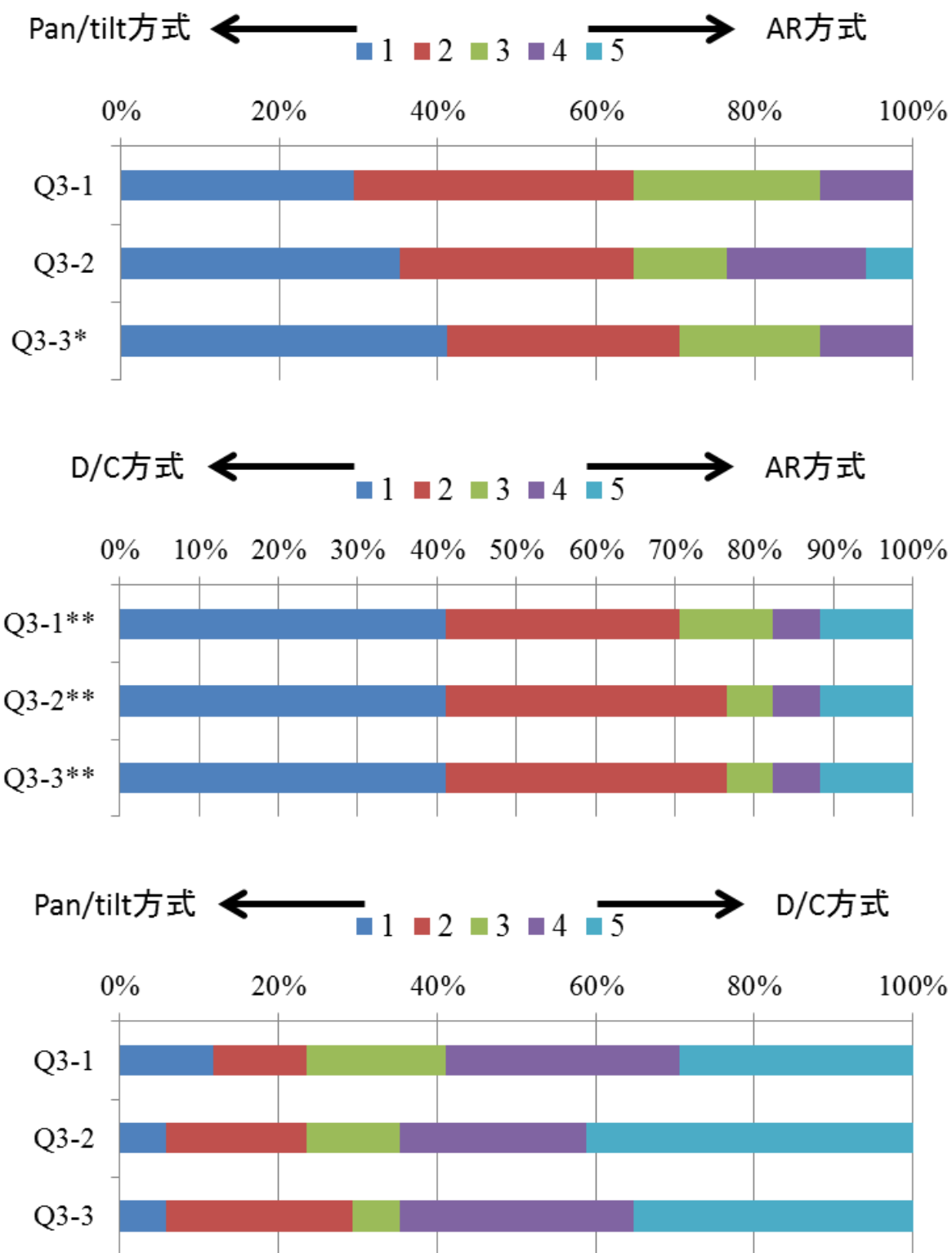


図 5-8 整列性を維持した視点操作方式の比較 主観的な印象評価

表 5-3 整列性を維持した視点操作方式の比較 タスク所要時間の平均

実験場所	AR 方式	Pan/tilt 方式	D/C 方式
A 地点	13.6 秒(5.9 秒)	13.7 秒(7.1 秒)	12.2 秒(6.7 秒)

()内は標準偏差.

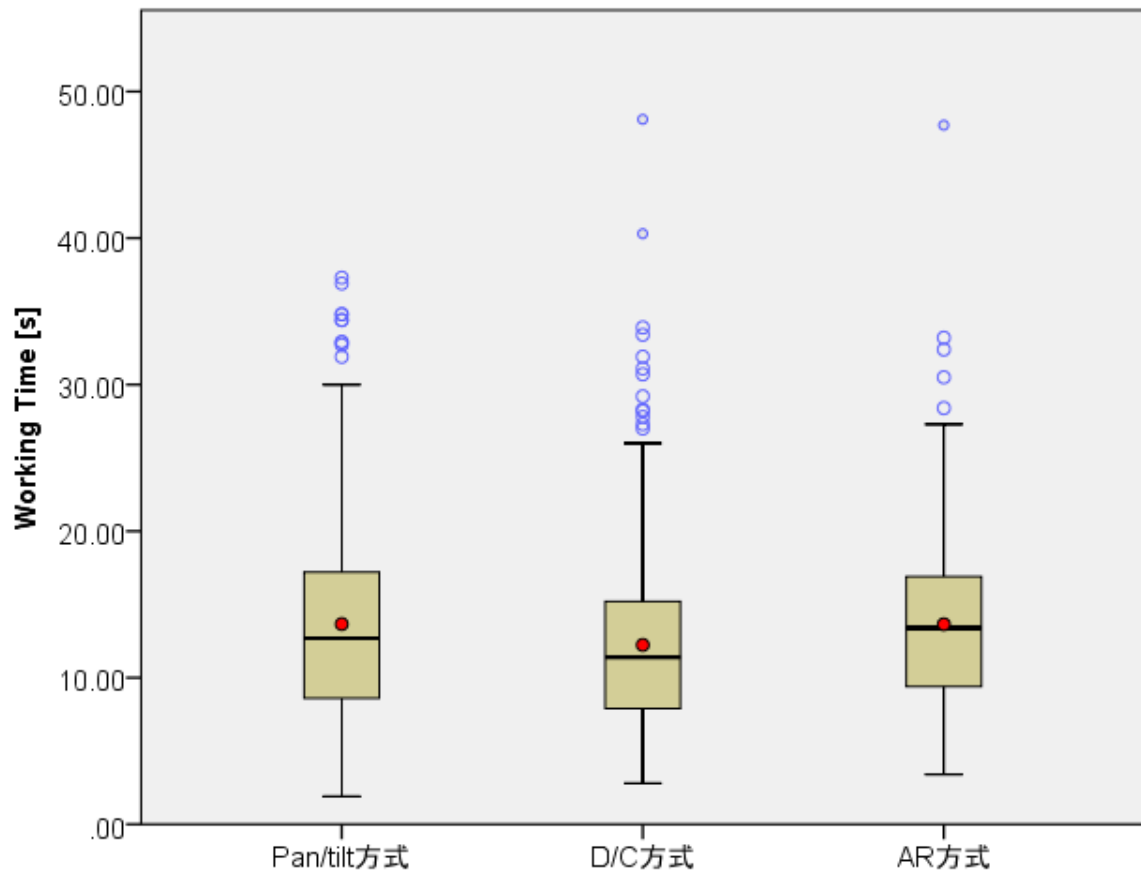


図 5-9 整列性を維持した視点操作方式の比較 タスクの所要時間

5.3.5 考察

実験の結果, D/C 方式は AR 方式に比べ, タスクの所要時間および主観的な印象評価の面で優れていた.

AR 方式は, 台の上においた建築模型と同様の見え方をするというメタファがあるにも関わらず, 印象評価においても良い評価が得られなかった. この原因の一つとして, AR 方式の実装に利用した AR マーカによるモバイル端末のカメラの位置姿勢推定が不安定になる場合や, マーカ

がカメラの撮影範囲からはみ出した場合に、失敗する場合があるという実装上の問題が挙げられた。一方、6名の被験者が「初めはAR方式が使いやすかったが、慣れるとPan/tilt方式やD/C方式が使いやすかった」とコメントした。また、3名の被験者がARマーカの周辺を動きまわるのが煩わしいとコメントするなど、位置姿勢推定の問題以外の原因もみられた。

Pan/tilt方式は、印象評価においてはAR方式よりも好まれたが、タスクの所要時間に有意な差は見られなかった。タッチスクリーンを用いて視点を平行移動させるのに必要な時間は、ARマーカの周囲を動きまわるのに必要な時間と大きな差が無かった。この実験の設定では、ARマーカの周囲を数歩で移動できるため、動きまわるのに長い時間は必要としない。ただし、位置姿勢推定に失敗した場合は、マーカの写り方を微調整するための時間も加わる。

D/C方式の印象評価においては、12人（全体の7割以上）の被験者がAR方式よりも優れていると回答した。さらに、11人の被験者がPan/tilt方式よりも優れていると回答した。D/C方式はジオラマモデルが常に画面内に提示されるため、ジオラマモデルとその上の矢印アイコンを観察しやすい。ジオラマモデルが観察しやすかったために、ユーザの見ている現実空間の風景と見比べやすく、結果として指示位置を理解しやすくなった可能性がある。

D/C方式の課題として、画面中に占めるジオラマモデルの表示面積の割合が大きいため、実写映像の大部分がジオラマモデルに隠され、写真撮影タスクを遂行する際にカメラを狙った場所に向けることが難しくなることが挙げられる。実写映像の縮小表示を利用してタスクの遂行は可能であるが、被験者からは「ジオラマモデルの表示／非表示を切り替えたい」というコメントも得られた。

この実験においては、ジオラマモデル上の矢印アイコンが建物のモデルに隠されるという状況がしばしば生じたが、建物のモデルの半透明表示（図 3-3a）によって矢印が観察できるので、被験者からのコメントはなく、実験中では大きな問題にはならなかったと考えられる。

Pan/tilt方式とD/C方式の比較では、印象評価においては有意な差は見られなかった。この2方式の差を確かめるため、実験場所を追加して被験者実験を行った。

5.4 Pan/tilt方式とDolly-round/crane方式の比較実験

Pan/tilt方式とDolly-round/crane方式の違いについて、A地点とB地点で実験を行った結果を示す。A地点は三方が建物で囲われた建物の密集する地点で、ジオラマモデル上の矢印アイコンが建物のモデルに隠されるという状況が発生しやすいのに対し、B地点は周囲の建物まで十～数十メートルの距離がある、比較的開けた地点である。B地点では、図 5-12 に示す20箇所を指示位置として利用した。

5.4.1 A 地点の実験結果

図 3-5 中の A 地点で行った実験の結果を示す。被験者は当研究室の 20 歳代の学生 13 名（男性 12 名女性 1 名）である。図 5-10 は各設問に対して、1～5 の評点を付けた人数の割合である。Q3-1 で 9 割、その他の問いでも 6 割以上のユーザが D/C 方式の方が優れていると答えた。Wilcoxon の符号付き順位和検定の結果、各質問の有意確率 p は、Q3-1: $p=0.000$ と Q3-3: $p=0.016$ で有意差（有意水準 5%）があり、Q3-2: $p=0.446$ は有意差なしという結果となった。1 回のタスクの所要時間は表 5-4 に示すとおりであり、有意差は見られなかった。表 5-5 にタッチパネルの操作によるジオラマモデルの拡大／縮小、および視点の水平移動を利用した被験者数と平均タスク数を示す。拡大／縮小を利用した被験者数には両方式とも違いは見られなかったが、視点の水平移動を利用した被験者数は Pan/tilt 方式で 8 名、D/C 方式で 3 名となり、D/C 方式の方が少ない傾向が見られた。

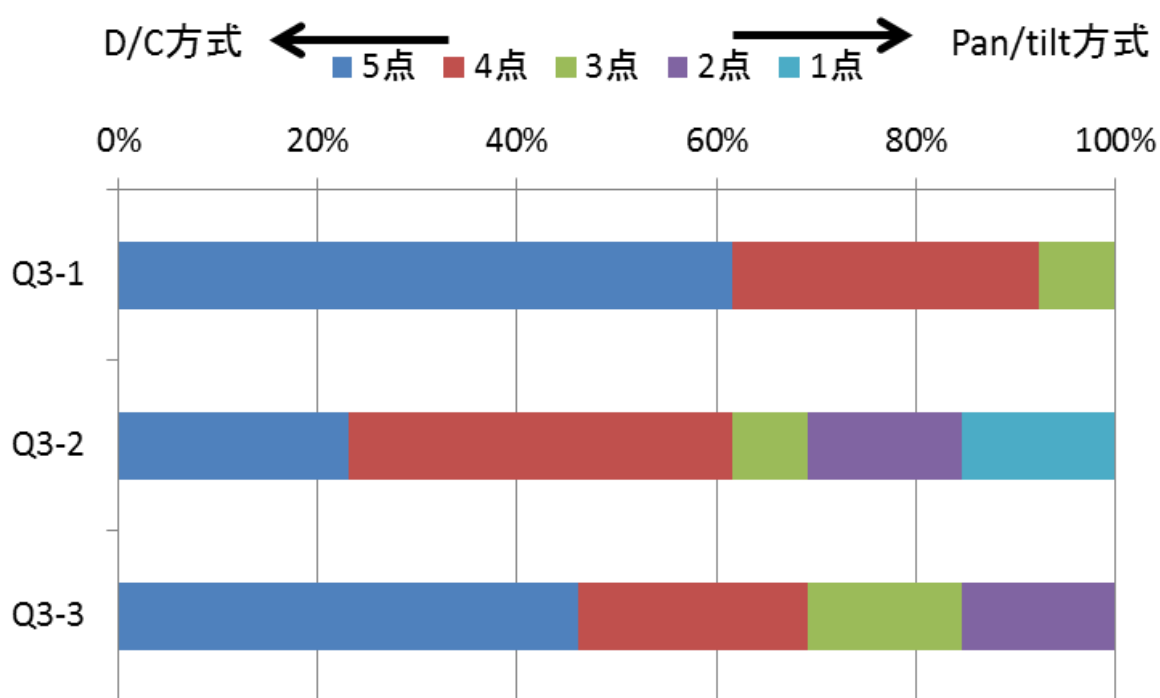


図 5-10 Pan/tilt 方式と D/C 方式の比較実験 A 地点 印象評価結果

表 5-4 Pan/tilt 方式と Dolly-round/crane 方式の比較実験 タスク所要時間の平均

実験場所	Pan/tilt 方式	D/C 方式	p 値(対応のある t 検定)
A 地点	8.7 秒 (1.5 秒)	7.9 秒 (2.0 秒)	$p=0.357$
B 地点	16.2 秒 (6.2 秒)	12.5 秒 (3.5 秒)	$p=0.032$

0内は標準偏差

表 5-5 Pan/tilt 方式と D/C 方式の比較実験 A 地点 タッチパネル操作の利用状況

利用状況	拡大／縮小	視点の水平移動
使用せず	5 名	4 名
Pan/tilt のみ	3 名 (2.0)	5 名 (3.4)
D/C のみ	3 名 (3.0)	0 名
両方で使用	1 名 (Free:3.0, D/C:1.0)	3 名 (Free:6.0, D/C:3.3)

()内は 10 タスク中の平均利用タスク数

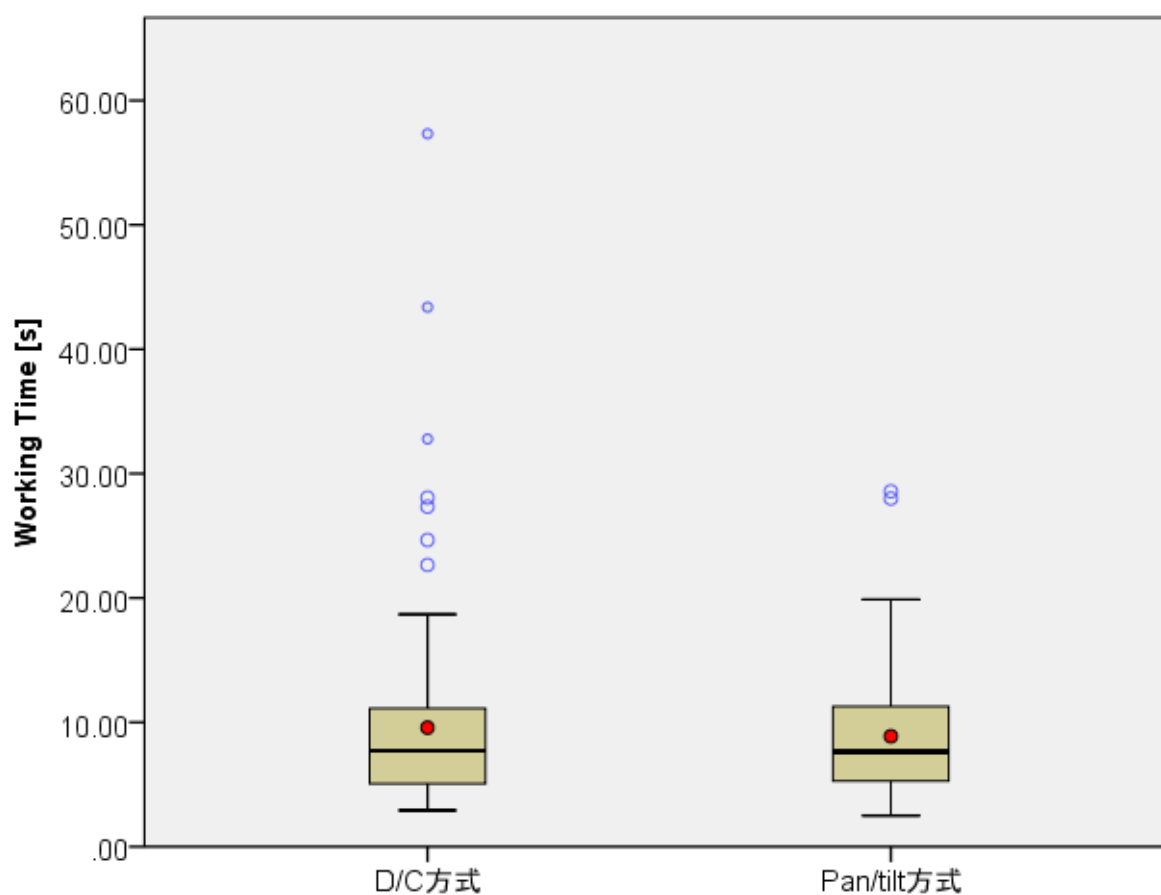


図 5-11 Pan/tilt 方式と D/C 方式の比較実験 A 地点 タスク所要時間

5.4.2 B 地点の実験結果

図 3-5 中の B 地点で行った実験の結果を示す。被験者は 20 歳代の男子学生 10 名である。被験者には指示位置がどこを指しているかわからなくとも 1 分間は諦めずに指示位置を探すように指示した。図 5-13 は各設問に対して、1～5 の評点を付けた人数の割合である。Q3-1 で 7 割、その他の問いでは 8 割のユーザが D/C 方式の方が良いと答えた。Wilcoxon の符号付き順位和検定の結果、Q3-2: $p=0.019$ と Q3-3: $p=0.023$ で有意差（有意水準 5%）があり、Q3-1: $p=0.277$ は有意差なしという結果となった。表 3-2 に示すとおり、1 回のタスクの所要時間は D/C 方式の方が有意に優れていた（有意水準 5%）。表 5-6 にタッチパネルの操作によるジオラマモデルの拡大／縮小、およびアンカーポイントの水平移動を利用した被験者数と平均タスク数を示す。拡大／縮小を利用した被験者数は両方式とも違いは見られなかった。視点の水平移動を利用した被験者数は Pan/tilt 方式で 9 名、D/C 方式で 6 名であり、Pan/tilt 方式の場合、10 タスク中平均 9 タスクで利用していた一方で、D/C 方式の場合は 10 タスク中平均 2.3 タスクでしか利用していなかった。

●先に行う10タスク △後に行う10タスク （番号は提示順序）

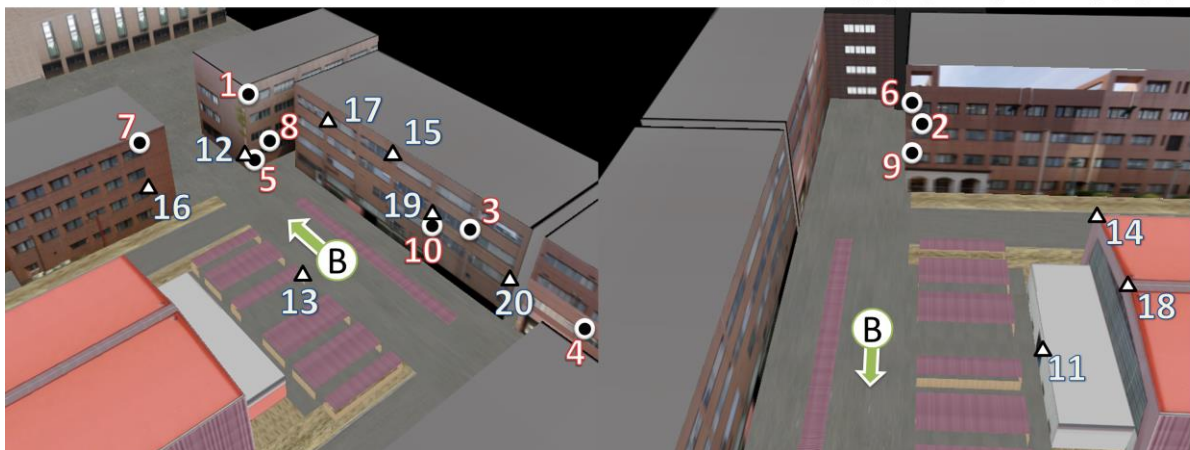


図 5-12 B 地点で行う写真撮影タスクに用いる指示位置

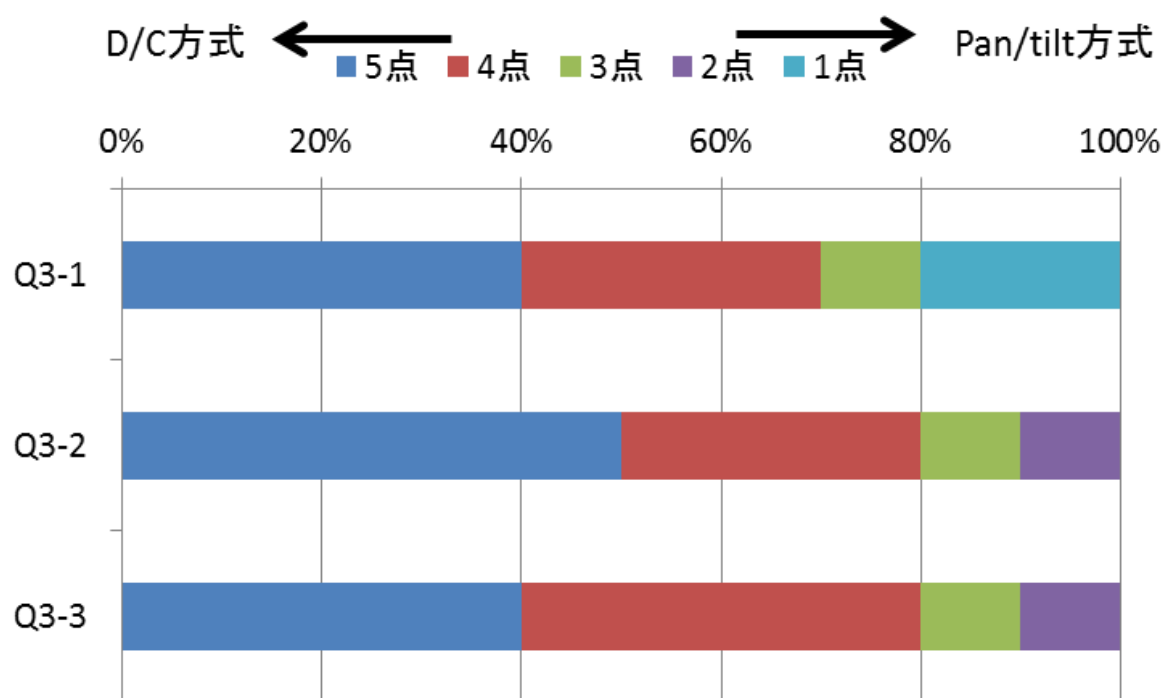


図 5-13 Pan/tilt 方式と D/C 方式の比較実験 B 地点 印象評価結果

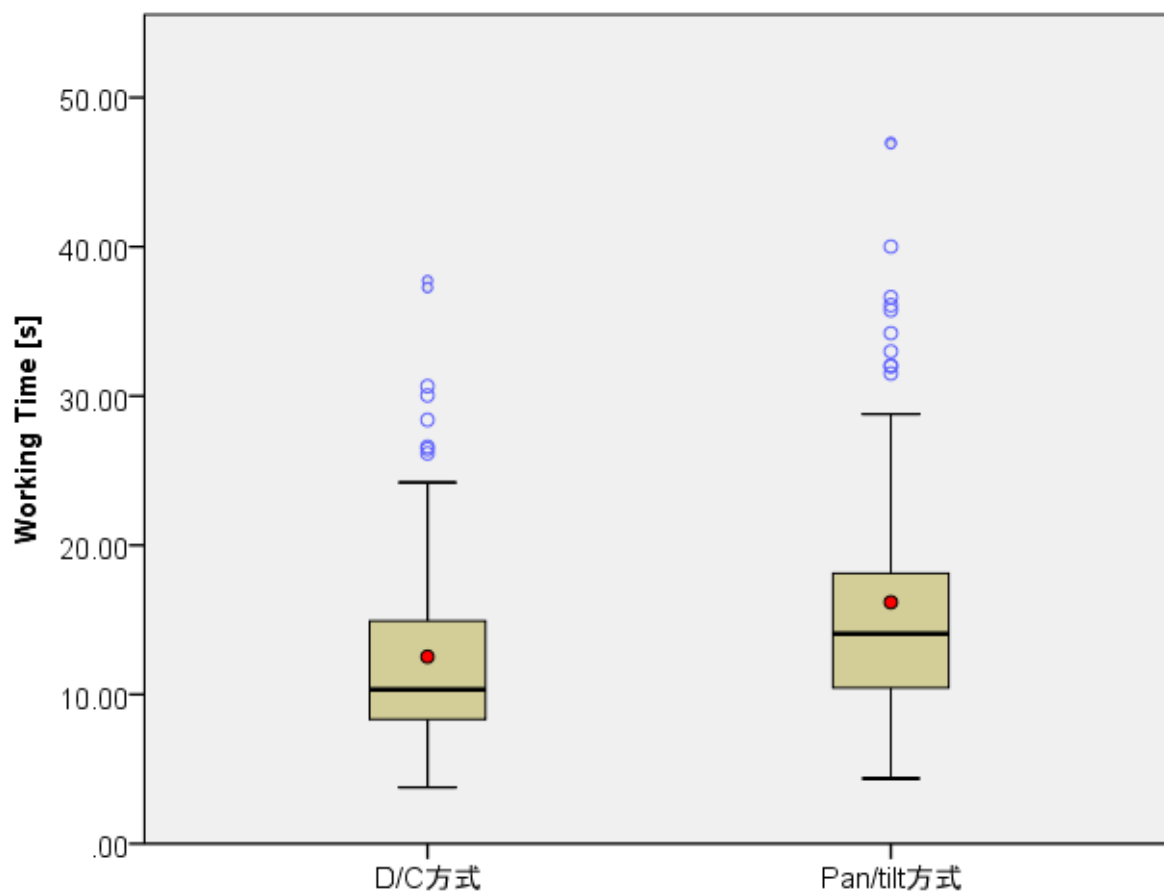


図 5-14 Pan/tilt 方式と D/C 方式の比較実験 B 地点 タスク所要時間

表 5-6 Pan/tilt 方式と D/C 方式の比較実験 B 地点 タッチパネル操作の利用状況

利用状況	拡大／縮小	視点の水平移動
使用せず	2 名	1 名
Free shot のみ	1 名 (2.0)	3 名 (8.0)
D/C shot のみ	1 名 (1.0)	0 名
両方で使用	6 名 (Free:4.5, D/C:3.3)	6 名 (Free:9.5, D/C:2.3)

()内は 10 タスク中の平均利用タスク数

5.4.3 考察

Pan/tilt 方式と D/C 方式の比較実験の結果, 印象評価については, A 地点, B 地点とも, Pan/tilt 方式の方が印象評価に優ると答えた被験者もいたものの, 多くは D/C 方式の方が優れていると評価した. D/C 方式の方が被験者らに好印象で受け入れられたと言える. 一方, 1 回のタスクの所

要時間を見ると、A 地点では有意差はないものの D/C 方式の方が短い傾向にあり、B 地点では D/C 方式の方が有意に短かった（有意水準 5%）。A 地点の実験で有意差が見られなかったことから、所要時間の面では不利と予想された Pan/tilt 方式も、状況によっては D/C 方式と大差ないという興味深い結果が得られた。B 地点では、Pan/tilt 方式、D/C 方式とも、1 回のタスクの所要時間が A 地点よりも長かった。これは、B 地点の方は建物が多く、指示位置の把握が難しい環境であるためと考えられるが、D/C 方式は Pan/tilt 方式に比べて時間の増加を抑えていることが確認できる。これらの結果より、指示位置の示す位置がわかりやすく、ジオラマモデル配置を操作しやすい方式は D/C 方式であることが明らかである。

タッチパネル操作を見ると、Pan/tilt 方式の方が視点の水平移動の利用者が多かった。しかし、全く使用しなかった被験者もいた。視点を水平移動させずとも指示位置の示す位置を理解できた原因として、今回の被験者らが実験環境をよく知っていたため、図 3-3(a)のような透過処理だけで十分に理解できた可能性がある。

地点別に見ると、A 地点より B 地点の方が視点の水平移動の利用が多く、Pan/tilt 方式においては 10 名中 9 名の被験者が 10 回中 8～9 回のタスクで利用していた。B 地点は、ジオラマモデル上に表示した指示位置と建物とが重なりやすい場所であり、指示位置の示す位置が透過させて表示した建物にあるのか、またはその奥の建物にあるのかが固定視点からでは判断できなかったために移動が必要となった可能性があり、実際にそのようなコメントをした被験者がいた。

第6章 直接的な指示共有方式との比較実験

複合現実型ジオラマインタフェースを用いた間接的な指示共有方式と、AR アノテーションによる直接的な指示共有方式との比較を行うために、実際の屋外作業を想定した、複雑なタスクを設定し、被験者実験を行う。

ユーザが屋外環境中のどこに居ても指示共有できるシステムを実現するためには、モバイルカメラの位置姿勢をリアルタイムに推定する必要がある。そのために、本論文では、屋外におけるビジョンベースのカメラトラッキング手法についても検討する。

6.1 AR アノテーションインタフェース

指示共有を実現する一般的な手法として、指差しやレーザポインタによって指示したい対象を直接指し示す手法がある。これらの手法による指示共有を、遠隔地にいる人々と実現するために、ビデオ会議システム [9] [10] [11]や、遠隔操作可能なレーザポインタを使ったシステム [12] [42]が提案されてきた。しかし、これらのシステムは、専用のデバイスが必要であり、カメラ付きのモバイル端末のみを用いて実現することが難しい。カメラ付きモバイル端末のみを用いて実現できる直接的な指示共有手法として、AR 技術を用いてアノテーション提示を行う方式がある。

本論文では、複合現実型ジオラマインタフェースの有効性を評価するため、比較対照として AR アノテーションインタフェースを実装する。AR アノテーションインタフェースでは、ビデオシースルー型の拡張現実映像を利用する。図 6-1 に示すように、モバイル端末のカメラで撮影された映像上に矢印などのアノテーションを直接重畳する。

AR アノテーションインタフェースの場合、仮想空間の位置合わせに誤差が生じると矢印の提示位置がずれてしまい、指示位置が誤って伝わる恐れがある。そこで、本論文では、ユーザが位置合わせのずれ具合を把握して、アノテーションの正しい位置を推測できるように、建物の CG モデルの輪郭線（輪郭モデル）を描画する。これにより、輪郭モデルが実際の建物とずれている場合でも、ユーザがその分を補正して理解することができる。また、位置合わせの精度が回復するように、モバイル端末をゆっくり動かすなどの対応が可能となる。

指示位置を入力する際は、望む位置でタップすると、その点からのレイキャスト [5]を用いて実際の建物に位置合わせした形状モデルとの交点を計算し、その交点が指示位置として入力される。

AR 型のアノテーションで位置情報を提示する場合、CG の位置合わせが正確にできていたとしても、アノテーションの奥行きを知覚することが難しいという問題がある。この問題は、仮想的な奥行きの手がかりを表示することで改善することが確認されているため、様々な映像表現が提案されている [4] [6]。このことを踏まえて、本章の実験で用いる AR アノテーションインタフェースには、直接的な指示位置アノテーション提示（図 6-1 中の青い矢印）に加えて、いくつかの増強を行う。まず図 6-1 に示すように、直接的な指示として、ビデオシースルー型の AR アノテーション表示（実写映像上への矢印の重畳）による指示位置の提示を行う。さらに、本実験で被験者に課す屋外カメラ設置におけるポーズ決定タスクを実行するために、屋外カメラの撮影範囲を視体積によって可視化する。加えて、奥行きをより知覚しやすくするため、屋外カメラのポーズからジオラマモデルを見た際の仮想映像も提示する（図 6-8 サブウィンドウ部）。この仮想映像は、指示位置を観察する第二の視点として利用できるため、ユーザの現在地から観察した指示位置の情報と合わせることで、指示位置の奥行きを正しく把握することができる。

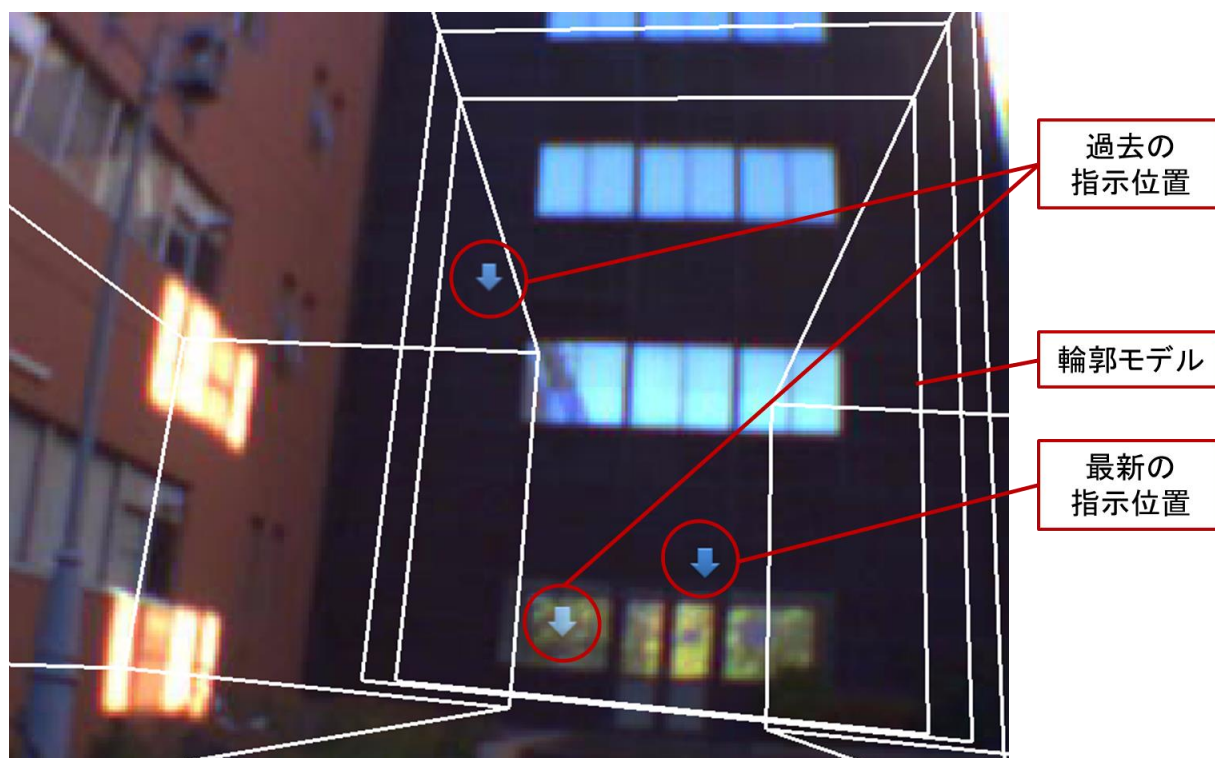


図 6-1 AR アノテーションインタフェース

6.2 撮影範囲の可視化

環境中のカメラの設置位置および撮影範囲をAR/MR技術を用いて可視化する取り組みとして、Kameda ら [43]による視体積の可視化や、住谷ら [44]による仮想浮動ミラーがある。環境カメラの撮影範囲を可視化する場合、視点が撮影範囲の内側にある場合と外側にある場合とで、ユーザーにとってわかりやすい方式は異なると考えられる。複合現実型ジオラマインタフェースでは、ジオラマモデル上で撮影範囲を可視化するため、ジオラマモデルを観察する仮想視点は撮影範囲の外側にあり、Kameda ら [43]の提案した視体積による可視化が有効であると考えられる。

[43]では、視体積を四角錐として可視化し、ビデオシースルー型のAR技術を用いて、実写映像上に重畳表示している。本論文では、図 6-2 に示す4種類の可視化方式（Volume, Arrow, Animation, Arrow+Volume）を挙げ、これらをジオラマモデル上で表示した場合と、ビデオシースルー型のARアノテーションとして表示した場合とで、撮影範囲のわかりやすさを評価した。Volume方式（[43]ではSurface method）は、四角錐の各面を半透明の色で表示する。Arrow方式は、四角錐と地面とが交わる面上に、カメラの位置に向かう半直線（Arrow）を一定間隔で表示する。Animation方式は、時間経過につれて徐々に大きくなる四角錐の各面を半透明の色で表示する方法である。四角錐の大きさが地面と交差するサイズ（Volume方式と同じサイズ）になると、大きさをゼロにするというアニメーションを1秒間隔で繰り返す。Arrow+Volume方式は、Volume方式とArrow方式を組み合わせた方式である。

20代の被験者9名に、ジオラマモデル上で視体積を可視化したビデオ映像と、ARアノテーションとして視体積を可視化したビデオ映像を提示し、撮影範囲がわかりやすい順に1～4番目まで順位を付けさせた。

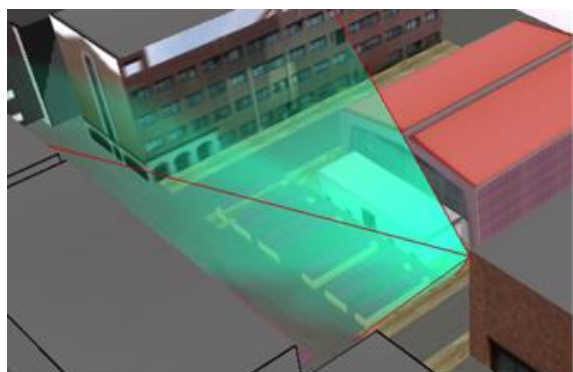
その結果、ジオラマモデル上でカメラの撮影範囲を表示した場合、図 6-3 に示すとおり、Animation方式もしくはArrow+Volume方式が最もわかりやすいと答えた被験者はともに4名で意見がわかれた。また、Volume方式が2番目にわかりやすいと答えた被験者は6名であった。

ARアノテーションでカメラの撮影範囲を表示した場合、図 6-4 に示すとおり、9名中6名はAnimation方式が最もわかりやすいと答えた。Arrow+Volume方式が1番目もしくは2番目にわかりやすいと答えた被験者がそれぞれ3名であり、被験者らの意見がわかれるという、ジオラマモデル上で表示した場合とは異なる傾向が得られた。

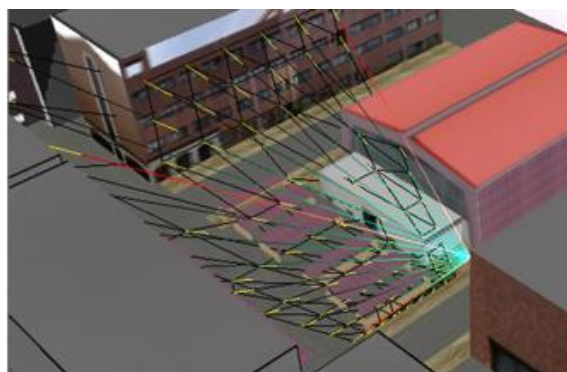
ジオラマモデルとARアノテーション型で共通の視体積の可視化方式とする場合には、どちらのビデオ映像においても1番わかりやすいという回答の多かった、Animation方式が有効であると考えられる。

いずれのビデオ映像においても、Arrow方式のわかりやすさは4番目と評価する被験者が多かった。その理由の一つとして、この実験で用いたビデオ映像は、視点が視体積の外にあることが挙げられる。Arrow方式は、特に視点が視体積の内側にあるときに地面と視体積が交わる範囲をはっきり示すことに配慮した方法であるため、視点が視体積の外にあるときには、地面と交わ

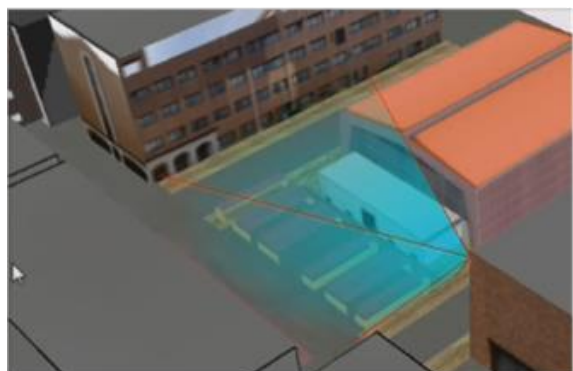
る範囲だけでなく、視体積の形状がはっきりわかるその他の方式の方が好まれたと考えられる.



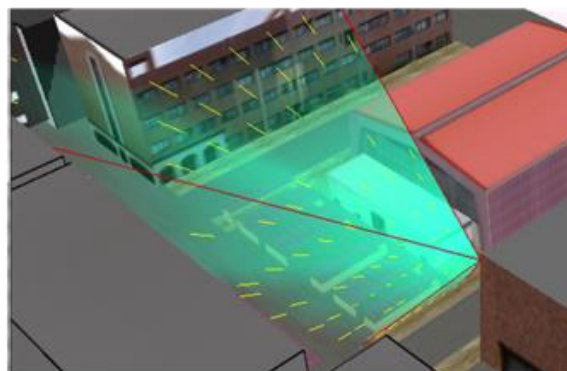
Volume



Arrow



Animation



Volume+Arrow

図 6-2 撮影範囲を可視化するための視体積の表示方式

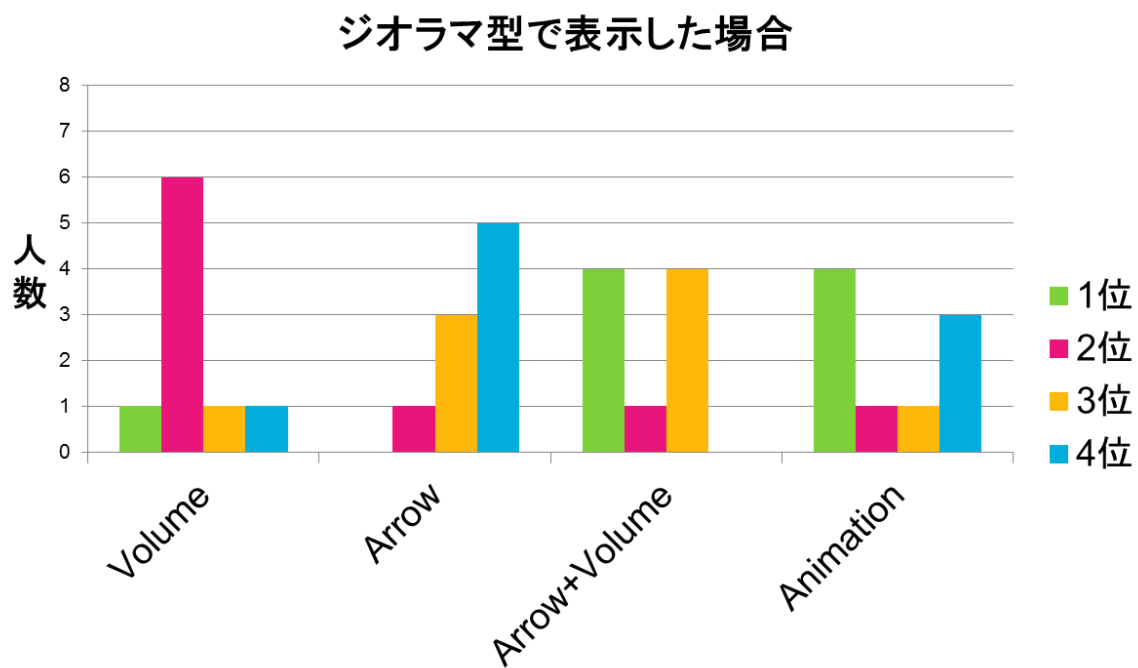


図 6-3 ジオラマモデル上でカメラの撮影範囲を表示した場合のわかりやすさ

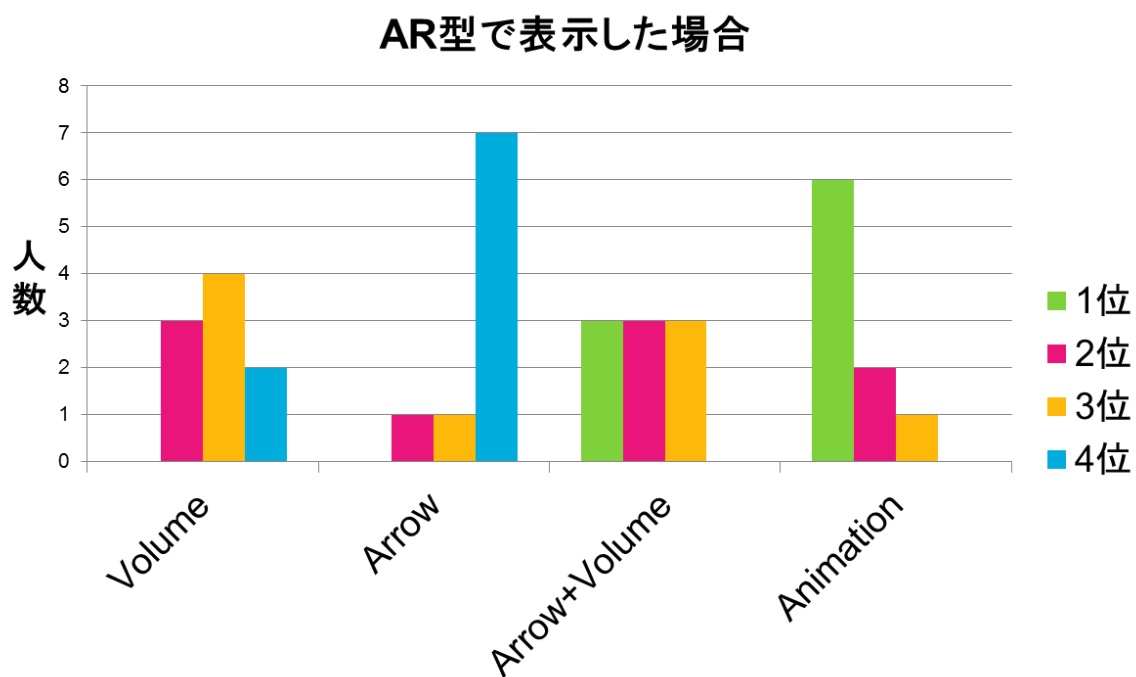


図 6-4 AR アノテーション型でカメラの撮影範囲を表示した場合のわかりやすさ

6.3 モバイル端末の位置姿勢推定

ユーザが屋外環境中のどこに居ても指示共有できるシステムを実現するためには、モバイルカメラの位置姿勢をリアルタイムに推定する必要がある。そのために、本論文では、屋外におけるビジョンベースのカメラキャリブレーション手法について検討する。

6.3.1 ビジョンベースの位置姿勢推定手法

屋外における複合現実感を実現するために、ビジョンベースのカメラキャリブレーション手法が提案されている [45] [46]。

代表的なアプローチの一つに SLAM (Simultaneous Localization and Mapping) 技術がある。特に、Klein ら [47] の提案した PTAM (Parallel Tracking and Mapping) は、未知の環境における特徴点の 3 次元マップ構築と、構築したマップに基づくカメラトラッキングを並列処理することで、比較的高速に実行可能で高いトラッキング性能を持つ実装手法である。しかし、この手法ではマップのサイズに上限があるため、適用範囲は机の周辺などの狭い環境に限られていた。これを広域環境で利用するために、Castle ら [48] は、あらかじめ複数のマップを構築しておき、適切なマップを選択してトラッキングを行うマルチマップアプローチによるトラッキングを実現する PTAMM を提案している。本論文では、PTAMM を用いてカメラ位置姿勢の推定を行う。まず、シーンのマップを構築し、そのマップが基準とする座標系から、建物の形状モデルが基準とする座標系への変換行列を手動の位置合わせによって与える。

6.3.2 マップの事前構築手法による PTAMM の初期化処理

本論文では、PTAMM を用いて屋外環境の 3 次元の特徴点マップを事前に構築して利用する。PTAMM では、マップは復元時のカメラの初期位置・姿勢から決まる相対座標系に基づいており、マップ毎に異なる座標系（マップ座標系）を基準として復元されている。そのため、単一の世界座標系上で指定されている指示位置を表示するためには、世界座標系とマップ座標系の変換行列を求める必要がある。

その方法の一つに、マップの特徴点群から代表点を数点選び出し、測量器などで計測した世界座標系におけるその点の正確な座標値を与えることで、マップ座標系から世界座標系へのスケール付きユークリッド変換行列を求める方法がある [49]。しかし、PTAMM が構築する特徴点マップは、3 次元位置の誤差の大きい点が含まれているため、その影響を取り除いたスケール付きユークリッド変換行列を求めることが難しい。

本論文では、より簡便な方法として、マップの構築する際に、環境中に世界座標系における位置姿勢が既知の AR マーカ [41] を設置することで、マップ座標系から AR マーカの座標系へのスケール付きユークリッド変換行列を求める方法を用いる。この変換行列を記録しておくことで、

AR マーカを取り除いた後も，PTAMM により AR マーカの世界座標系を用いたカメラトラッキングを継続できる．

実験環境の特徴点マップを構築するにあたり，屋外環境における PTAM によるカメラ位置姿勢推定の精度評価を行った（付録 A 参照）．屋外歩行者が手持ちしているカメラで撮影した映像シーケンスを用いてカメラの位置姿勢推定を行った結果，カメラが移動するにつれて，位置の推定誤差が大きくなるという結果が得られた．そこで，本実験では，図 6-5 に示すように，実験環境中の複数地点でマップを構築し，各マップの座標系を AR マーカと手動の位置合わせによって統一して利用する．

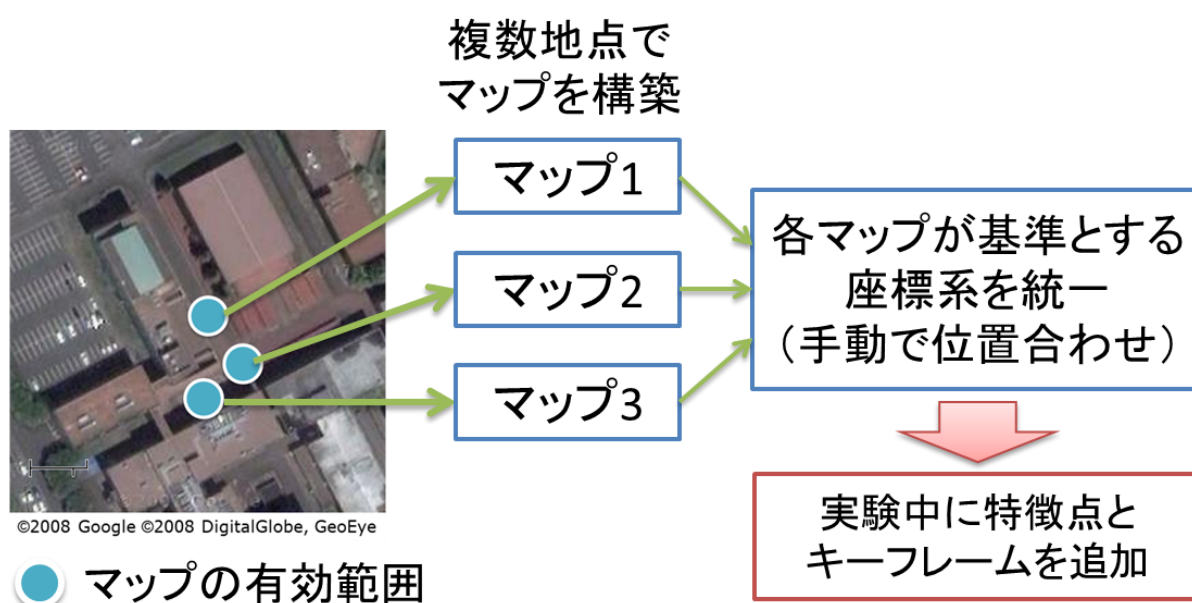


図 6-5 マップの事前構築による PTAMM の初期化

6.4 屋外カメラ設置におけるポーズ決定タスク

6.4.1 タスクの概要

屋外に監視・防犯等の目的で定点カメラを設置する際には，その設置場所を複数の人々が相談して決める必要がある．この状況を想定し，本実験では“監督者”と“作業者”の二人一組で指示共有を行う．

監督者が定点カメラで撮影したい画像のイメージを明確に持っており，作業者がそのイメージに沿って定点カメラの設置位置・角度を見つけ出す．監督者にはあらかじめ撮影したい画像を見せ

ておき、実験中にも随時参照させる。

手順として、まず監督者が定点カメラで撮影すべき位置を指示し、作業者がその位置を撮影できる定点カメラの候補位置・角度を推定し、その候補位置・角度を監督者に伝える。これを1サイクルとして、定点カメラの候補位置・角度が目標値と十分近くなった場合に作業を終了する。監督者と作業者が交わす言葉は、それぞれ「ここが写るように（または写らないように）」、「ここはどうですか？」のみに制限し、言葉による説明に頼らない指示共有について評価する。

本実験で利用する屋外カメラのポーズ決定タスクは、屋外にカメラを新設する際に、どのような位置・角度（ポーズ）で設置するかを複数の人々が相談して決める作業を想定している。

このタスクは、監督者と被験者の2人で行う。監督者は、実験を通して共通の人物が務める。監督者は位置の入力を行いやすくするため、モバイル端末姿勢の代わりにマウスによって操作するジオラマインタフェースを用いる。

監督者と被験者はハンズフリーのトランシーバを用いて会話を行えるが、タスクの実施中は言葉による説明で位置情報が伝わることを避けるため、位置情報を送信したことを言葉で伝える。タスク開始時、被験者は図 6-6 中に示す位置に立ち、タスクの実施中は必要に応じて移動する。

タスクの手順を以下に示す。

まず、監督者は指示共有インタフェースを用いて、被験者に屋外カメラで撮影すべき位置（被写体が将来来るであろう位置）を送信する。同時に、トランシーバで「ここが映るようにして下さい」という音声を被験者に伝える。

被験者は、その場所を撮影できる屋外カメラの位置・角度（ポーズ）を見つけ出す。その際、被験者がジオラマモデルや輪郭モデルのみを観察して実環境との対応を理解しないままポーズを決定することがないように、実環境中にのみ、屋外カメラの設置を禁止する場所を示す×マークを複数設置する。図 6-7 に示すよう

被験者には、屋外カメラの位置を決定する前に、必ず一度は実環境を観察し、×マークが存在していないことを確認するように指示する。

被験者は、適切な屋外カメラのポーズを見つけたら、指示共有インタフェースを用いて位置を入力し、角度を調整して監督者に送信する。同時に、「ここはどうですか？」という言葉をついでにトランシーバで伝える。

監督者は、受信した屋外カメラのポーズを確認し、他にも撮影すべき位置があれば、その位置を被験者に送信する。被験者は、新たに指示された位置も同時に撮影できるポーズを見つけ出し、再び監督者に送信する。

このようなやりとりを繰り返し、被験者の提示した屋外カメラのポーズで、監督者の要求する撮影すべき範囲を全て撮影することができる状態になればタスクを完了する。今回は、監督者の要求する撮影すべき範囲を明確にするため、目標のカメラ位置姿勢を決めておき、その視体積を監督者の見るジオラマモデル上に Arrow 方式で可視化する（図 6-2）。ジオラマモデルと視体積が交差する領域から 16 点を均等に抽出し、その全てを撮影できる屋外カメラのポーズを発見出来

れば，タスク完了とする．

タスク開始から 5 分以上経過してもポーズが目標と大きく違う場合には，タスク未完了として扱う．

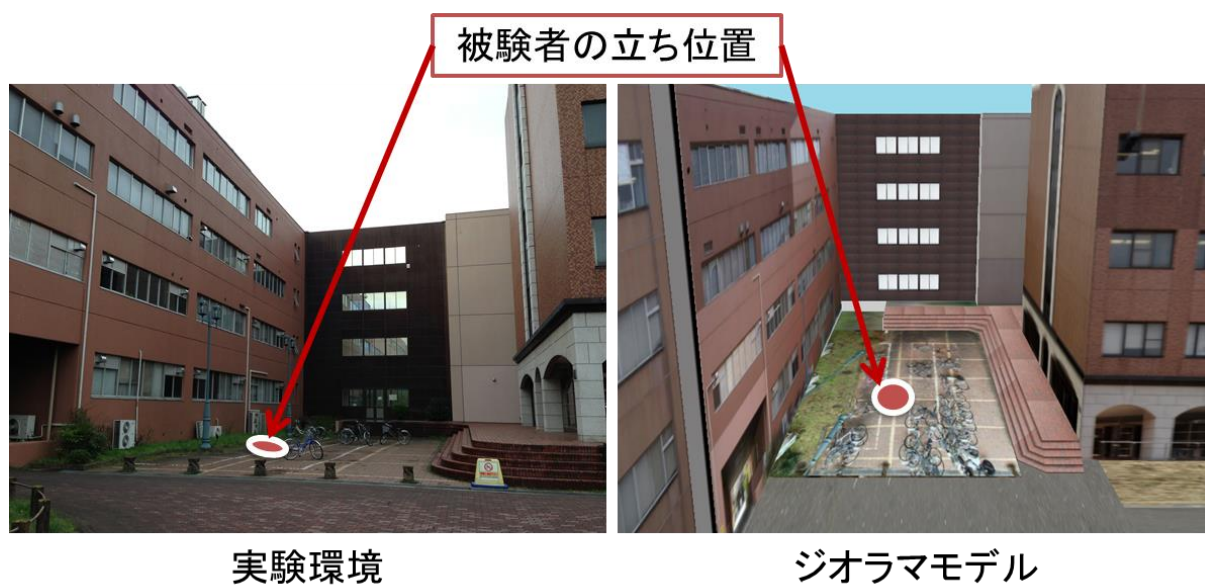


図 6-6 実験環境（左）とジオラマモデル（右）

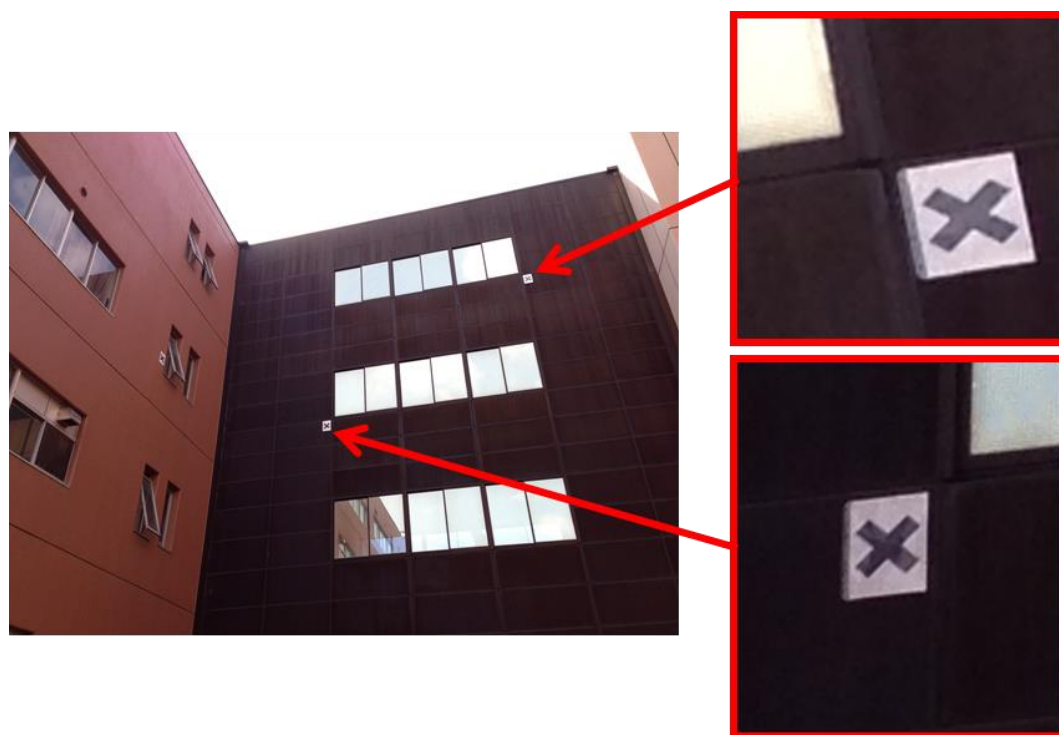


図 6-7 屋外カメラ設置禁止場所を示す×マーク

6.4.2 屋外カメラ設置におけるポーズ決定タスク用インタフェース

屋外に設置するカメラのポーズを検討するために、実験で用いる指示共有インタフェースに共通する機能として、屋外カメラ角度の入力と撮影範囲の可視化を導入する。

屋外カメラのポーズのうち、位置の入力は図 6-8 中のメインウィンドウ内のタップで行う。角度の入力は、右側のサブウィンドウ領域を用いる。サブウィンドウには、屋外カメラからジオラマモデルを観察している仮想映像を表示する。サブウィンドウ内で左右に指をドラッグすると屋外カメラが左右に回転する。上下にドラッグした場合は屋外カメラが上下に回転する。サブウィンドウをダブルタップすることで、メインウィンドウとサブウィンドウの役割を切り替えることができる。これにより、屋外カメラから見た仮想映像が大きく表示され、内容を観察しやすくなるとともに、ドラッグ操作による角度の微調整がしやすくなる。

このインタフェースでは、2 点の同時入力可能なマルチタッチスクリーンを用い、2 本指のピンチ操作でジオラマモデルの拡大／縮小を行う。

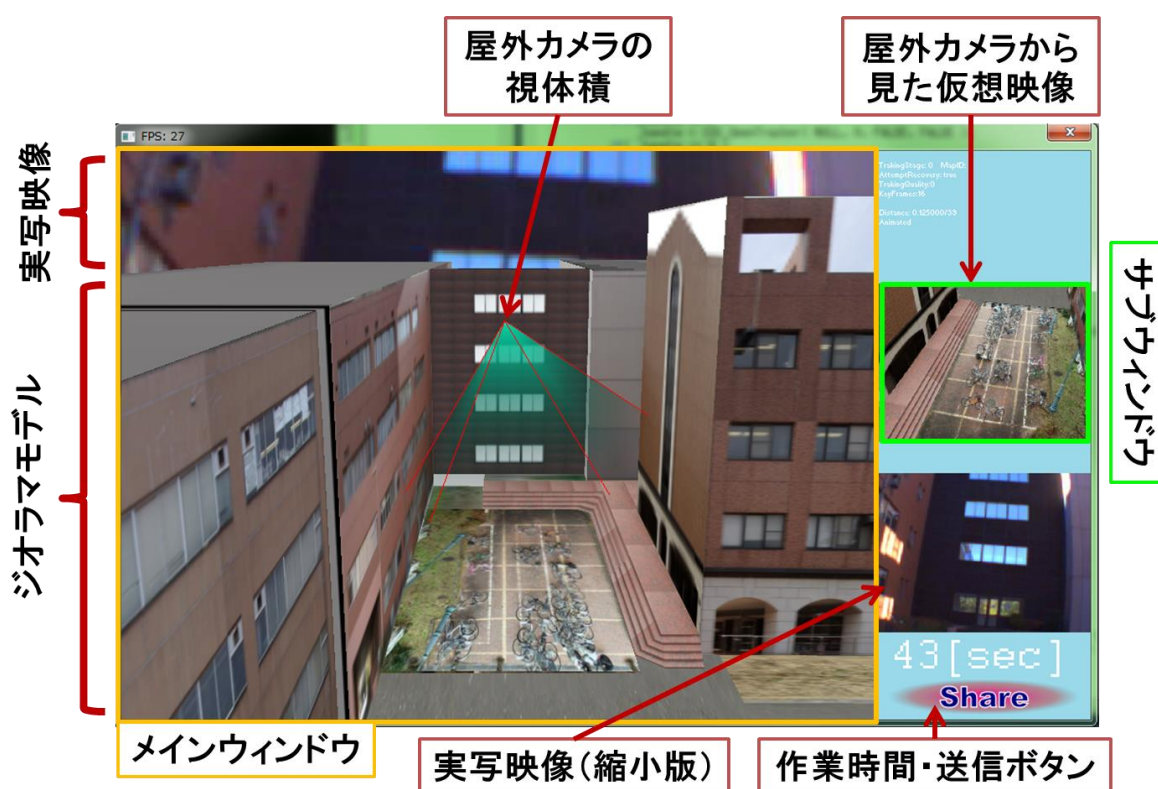


図 6-8 屋外カメラ設置におけるポーズ決定タスクのための画面構成

6.4.3 実験手順

被験者は、提案インタフェースと AR アノテーションインタフェースを順番に利用し、それぞれについて評価する。まず、どちらか一方の指示共有インタフェースについて操作方法の説明と

練習を実施した後、タスクを3回実施し、所要時間を計測した。その後、主観的な印象評価とコメントとを得た。次に、もう一方のインタフェースを用いて同様に、操作方法の説明と練習、3回のタスクの実施、主観的な印象評価とコメントの取得を行う。被験者らを半分ずつ、先に提案インタフェースを利用するグループと、先にARアノテーションインタフェースを利用するグループに分ける。

6.5 実験結果と考察

6.5.1 印象評価の結果と考察

主観的な印象評価の結果を示す。被験者は22～30歳の男性4名である。質問は表6-1に示す19個で、5段階のリッカート尺度を用いた。質問に同意できる場合を2、どちらとも言えない場合を0、質問に同意できない場合を-2とする。提案インタフェースをVD、ARアノテーションインタフェースをARとして、各質問に対する被験者の評点の平均と分散をグラフ化したものを図7に示す。このグラフより、Q6-10、Q6-11を除く、全ての質問で提案インタフェースが良い評価を得ていることがわかる。

Q6-1～Q6-3の結果からわかるように、提案インタフェースは指示された位置のわかりやすさ、入力しやすさだけでなく、両インタフェースで共通の操作を用いた屋外カメラの角度入力においてもよい評価であった。

また、Q6-5、Q6-6の結果から、ARアノテーションインタフェースの場合、被験者らは作業に難しさや疲れを感じていた。すべての被験者はコメントで、建物の上の方の位置を指示する際に、モバイル端末を上に向ける操作が身体的につらいと報告していた。

Q6-7～Q6-9の結果から、被験者らは提案インタフェースに対し、便利さや使いやすさを感じていたことがわかる。

Q6-10、Q6-11の結果より、ARアノテーションインタフェースにおいて、被験者がモバイル端末のカメラで撮影された実写映像を観察していたことがわかる。被験者らのコメントでは、遠い場所を観察する時は実写映像の解像度が不足したため、現実の風景を確認する必要があったことが報告されていた。

今回モバイル端末の位置姿勢推定に用いた、慣性センサとPTAMM [48]によるトラッキングは、日照条件の違いなどによって不安定になる場合があり、位置合わせに誤差が生じる場合がある。どちらの指示共有インタフェースも、この方式で位置合わせをしているが、Q6-17の結果の通り、ARアノテーションインタフェースでは位置合わせの安定性がより求められている。

表 6-1 直接的な指示共有方式との比較実験 主観的な印象に関する質問

Q6-1	指示された位置はわかりやすいか？
Q6-2	屋外カメラの位置は入力しやすいか？
Q6-3	屋外カメラの角度は、入力しやすいか？
Q6-4	適切な屋外カメラの位置を決めやすいか？
Q6-5	作業は難しいか？
Q6-6	作業は疲れないか？
Q6-7	位置情報を伝える道具として便利か？
Q6-8	この道具を日常生活の中で使いたいのか？
Q6-9	インタフェースは使いやすいか？
Q6-10	インタフェースの後ろの実写映像は参考になるか？
Q6-11	インタフェースの後ろの実写映像は見やすいか？
Q6-12	サムネイルの映像は参考になるか？
Q6-13	インタフェースに慣れるのは簡単か？
Q6-14	作業は楽しいか？
Q6-15	位置情報のコミュニケーションは円滑だと思うか？
Q6-16	会話が自由にできたら効率が上がると思うか？
Q6-17	カメラトラッキングの安定性に問題はないか？
Q6-18	画面に提示されていた情報は十分か？
Q6-19	初期位置から移動する必要性を感じるか？

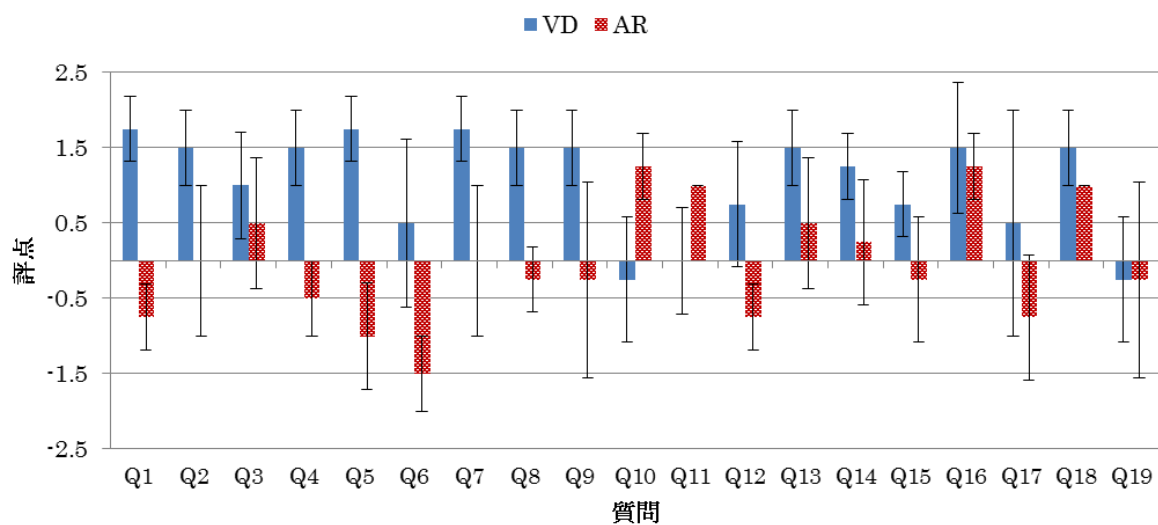


図 6-9 主観的な印象評価の結果

6.5.2 タスク所要時間の結果と考察

図 6-10 のグラフにタスクが完了するまでの所要時間と標準偏差を示す。提案インタフェースを用いた3回の試行をVD1～VD3, AR アノテーションインタフェースを用いた3回の試行をAR1～AR3 とする。被験者 1 と被験者 2 は, AR2 および AR3 において, 5 分以上かけてもタスクをで終えることができなかったため, 未完了として扱った。

図 6-11 はタスクが未完了の被験者を除いた, タスク毎の平均所要時間と標準偏差を示す。提案インタフェースを用いた場合と比較して, AR アノテーションインタフェースを用いた場合は所要時間が長くなる傾向がある。この実験では被験者の数が限られているため, 今後より多くの被験者で実験を実施して検証する必要がある。

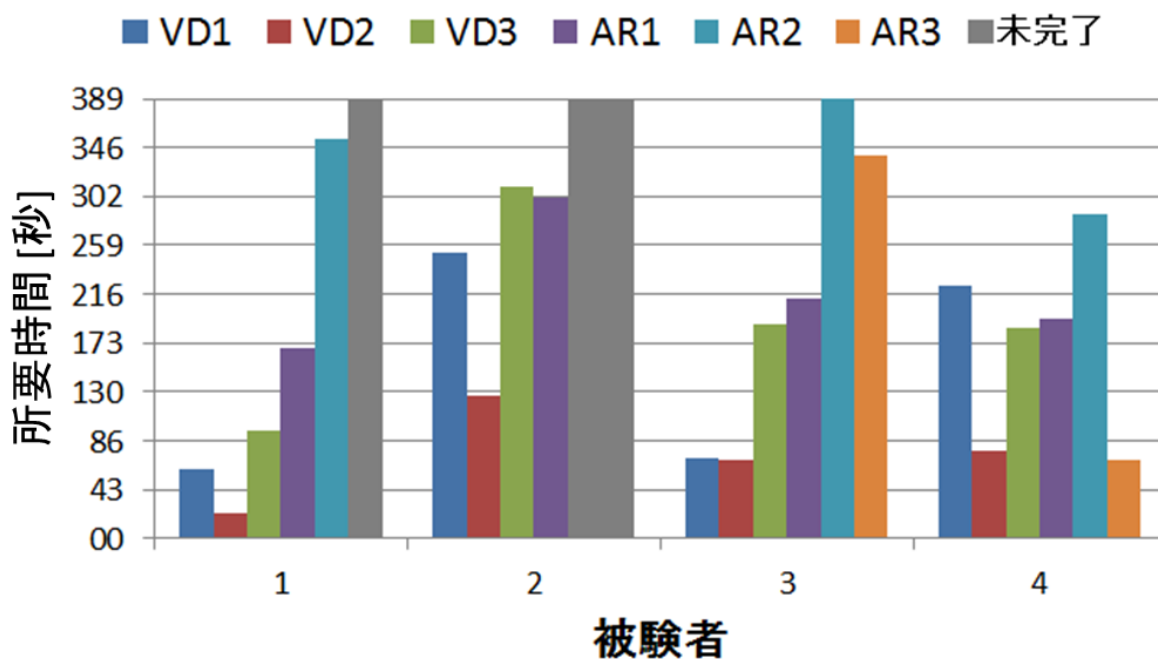


図 6-10 第 6 章 直接的な指示共有方式との比較実験 被験者毎の所要時間

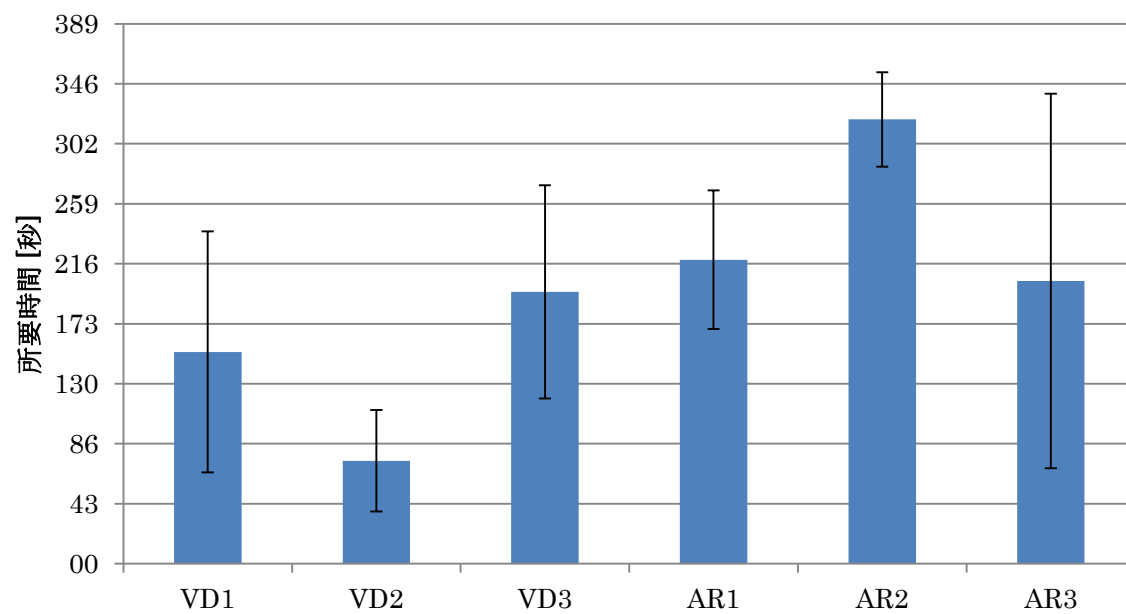


図 6-11 直接的な指示共有方式との比較実験 タスク毎の平均所要時間

第7章 結論

本論文では、屋外環境にいる人々が、各人の持ち場に居ながら互いに的確な3次元位置情報の指示共有を実現するために、複合現実型ジオラマインタフェースを提案した。提案インタフェースは、建物などの写実的な3次元CGモデル（ジオラマモデル）が、ユーザの目前に存在しているかのようにモバイル端末上に提示するインタフェースである。このジオラマモデル上で3次元位置情報の入力と提示を行う。提案インタフェースを用いて、入力した位置情報を、無線ネットワークを介して他のユーザが持つモバイル端末に送信し、提示することで、指示共有を実現する。

直接的な指示方法と比較して、(1) モバイル端末の位置姿勢推定の誤差の影響を受けにくい、(2) 指示位置がユーザの視野外にあっても、ジオラマモデルを観察する仮想視点を移動させるだけで指示共有が可能である、(3) 様々な視点から3次元位置を観察することで、正しい3次元位置を把握できる、という利点がある。ただし、ユーザが指示位置を理解するためには、ジオラマモデルと現実空間との対応を把握する必要がある。本論文では、この対応を把握するための、視覚的な手がかりを提示することにより、指示位置を理解しやすい複合現実型ジオラマインタフェースの設計について追求した。そのために実施した被験者実験の概要と結果を、以下にまとめる。

① ジオラマモデルを実写映像へ重畳する方式の是非の検証

写真撮影タスクを用いた被験者実験の結果、タスクの所要時間に有意な違いは見られなかった。印象評価の結果、被験者の意見はわかれており、実写映像への重畳することが指示位置をわかりやすさを改善していると考えている被験者は全体の半数程度であった。

被験者らのコメントから、実写映像へ重畳する際には、単純に実写映像とジオラマモデルを重ね合わせるだけでなく、実写映像とジオラマモデルの画面中に占める面積比などを考慮する、ジオラマモデルの境界線を明確に表示する、などの対策によって、重畳表示を指示位置のわかりやすさ向上に結び付けられるよう方式を改善できる可能性があると考えている。

② 整列性を維持した視点操作方式の比較

写真撮影タスクを用いた被験者実験の結果、ジオラマモデルと現実空間とを整列させる方式（整列方式）と、ジオラマモデルの見え方（仮想視点の向き）を固定する非整列方式とを比較した。タスクの所要時間に有意差は見られなかったが、印象評価においては、整列方式画面が見やすく、指示位置がわかりやすいという結果が得られた。

そこで、整列性を維持した視点操作方式として、Augmented-Reality (AR)方式、Pan/tilt 方式、

Dolly-round/crane (D/C) 方式の 3 種類を挙げ、写真撮影タスクによる被験者実験で、視点操作のしやすさと指示位置のわかりやすさについて比較を行った。その結果、AR 方式や Pan/tilt 方式に比べ、D/C 方式はタスク所要時間、主観的な印象評価ともに優っていた。

Pan/tilt 方式と D/C 方式については、実験地点を増やし、検証を重ねた。その結果、印象評価の面では、実験地点によらず D/C 方式が優れていた。タスクの所要時間については、実験地点によって異なり、作業環境の複雑さ（建物の数や建物までの距離など）の影響を受けると考えられるが、D/C 方式は Pan/tilt 方式に比べて所要時間の増加を抑えていることが確認できた。

③ 直接的な指示共有方式との比較

①、②の被験者実験の結果から、指示位置がわかりやすい複合現実型ジオラマインタフェースを実現するために、ジオラマモデルを実写映像へ重畳し、D/C 方式による仮想視点操作を行うこととした。このインタフェースを用いた指示共有の有効性を評価するため、AR アノテーションインタフェースとの比較を行った。

AR アノテーションインタフェースは、モバイル端末のカメラで撮影した実写映像に指示位置を示す矢印（アノテーション）を重畳することで、指示位置を直接指し示すインタフェースである。モバイル端末を用いて、遠隔地にいる人々が直接的な指示共有を行える手法である。

屋外環境における実用的な共同作業の例として、屋外カメラ設置におけるポーズ決定タスクを挙げ、アプリケーションを構築して実験を行った。その結果、本論文で提案する複合現実型ジオラマインタフェースは、AR アノテーションインタフェースに比べて主観的な印象評価とタスクの所要時間の面でよい評価を受けていることが確認できた。ただし、被験者の数が少いため、被験者を増やし、結果の信頼性を高める必要がある。

今後の課題

①の被験者実験では、ジオラマモデルを実写映像へ重畳することで、必ずしも指示位置がわかりやすくなるわけではないことが示された。被験者らのコメントなどから、実写映像とジオラマモデルの境界線をコントラストの高い色で強調表示してわかりやすくすることや、実写映像とジオラマモデルが画面中に占める面積比などに注意することで、重畳する場合を指示位置の見やすさ向上に結び付けられるよう方式を改善できる可能性が示されたため、検証が必要である。

本論文では、提案インタフェースを用いる以外の手段で指示共有がなされることを避けるため、音声によるコミュニケーションを制限して実験を行った。音声や、モバイル端末のディスプレイ側に取付けられている対面カメラから撮影したユーザの映像などを利用することで、指示共有のしやすさが改善できると考えられる。

応用例

提案インタフェースは、カメラや慣性センサなど、一般的なモバイル端末に搭載されているセンサ類を前提としている。既に、都市の写実的な3次元モデルを持った地図配信サービスも実用化されているため、これらを利用することで、スマートフォンやタブレット PC 向けに、様々な場所で利用可能な、複合現実型ジオラマイインタフェースを用いた指示共有アプリケーションが実現できる。

本論文では、屋外カメラ設置におけるポーズ決定タスクを例としてアプリケーションを設計し、被験者実験を行った。このアプリケーションでは、地面や壁面を撮影対象として指定し、そこが撮影されていれば屋外歩行者も撮影できるであろうと期待して、ポーズ決定を行った。これは、静止物体を指示対象とした指示共有といえる。

パレードやマラソンの中継撮影のように、移動する被写体を適切な角度から撮影するためのカメラワークを事前に計画するなど、移動物体を対象とした指示共有が必要な場面もある。提案インタフェースはこのような場面に対しても、ジオラマモデル上に移動物体の CG モデルを提示する、カメラを動かす方向と距離を矢印で指示共有できるようにする、などの機能追加を行うことで対応できると考えられる。

静止物体を対象とした指示共有が必要な共同作業の例として、建造物の劣化調査（ビルの壁面のヒビ割れ・破損箇所の指示共有）や、動植物の生息状況の協働調査（自治体が市民に協力を求め、動植物の生息地を調査する）などがある。移動物体を対象とした指示共有が必要な作業の例としては、パレードやマラソンの撮影、逃走する人や動物を複数の人々が協力して囲い込む、などの作業が挙げられる。いずれも、作業に応じてアプリケーションを構築しなければならないが、3次元位置情報の指示共有を行うための基礎的なインタフェースとして、複合現実型ジオラマイインタフェースの利用が期待できる。

謝辞

筑波大学 画像情報研究室に配属されてから 6 年以上にわたって懇切丁寧なご指導を賜りました，本学大学院システム情報工学研究科 大田友一教授，亀田能成准教授，および北原格准教授に，深く感謝いたします．ゼミやミーティング，論文執筆を始めとして様々な場面において，研究について，研究者としてのあり方について，厳しくも的確なご助言をいただきました．

本論文の査読をしてくださり，貴重な御意見・御指摘を賜りました，本学大学院システム情報工学研究科 葛岡英明教授，産業技術総合研究所 サービス工学研究センター 蔵田武志教授（連携大学院）に心から御礼申し上げます．

また，研究を進めるに当たり，数々の助言・議論いただきました本学画像情報研究室 OB・OG の皆様，特に配属以来，数々の御助言と手助けを下さいました，糟谷望先輩，Senya Polikovsky Semion 君，佐藤秀昭君に御礼を申し上げます．共同研究者として共に研究し，支えてくれた山崎真也先輩，吉田亮君，長島正典君に深く感謝いたします．研究を進めるにあたり，数々の助言・議論をいただき，様々な実験に快く参加・協力して下さった本学画像情報研究室の皆様に，心より感謝いたします．

参考文献

- [1] 孫 磊, 福田 知弘, 徳原 俊樹, 矢吹 信喜, "模型と VR との視点連携による都市プレゼンテーションシステムの開発--二つのマーカーによる精度向上と専門家による評価," *日本建築学会環境系論文集*, vol. 76, no. 668, pp. 953-961, 2011.
- [2] 古賀 元也, 鵜 心治, 多田村 克己, 大貝 彰, 松尾 学, "景観まちづくりにおける空間イメージ共有手法に関する研究," *日本建築学会計画系論文集*, vol. 73, no. 633, pp. 2409-2416, 2008.
- [3] Yuichi Ohta, Hideyuki Tamura, "Mixed Reality: Merging Real and Virtual Worlds," *Springer-Verlag New York*, 1999.
- [4] Tobias Höllerer, Steven Feiner, Tachio Terauchi, Gus Rashid, Drexel Hallaway, "Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System," *Computers & Graphics*, vol. 23, no. 6, pp. 779-785, 1999.
- [5] Jason Wither, Stephen DiVerdi, Tobias Höllerer, "Annotation in Outdoor Augmented Reality," *Computers & Graphics*, vol. 33, no. 6, pp. 679-689, 2009.
- [6] Arindam Dey, Andrew Cunningham, Christian Sandor, "Evaluating Depth Perception of Photorealistic Mixed Reality Visualizations for Occluded Objects in Outdoor Environments," *17th ACM Symposium on Virtual Reality Software and Technology*, pp. 211-218, 2010.
- [7] 小田島 太郎, 神原 誠之, 横矢 直和, "拡張現実感技術を用いた屋外型ウェアラブル注釈提示システム," *画像電子学会誌*, vol. 32, no. 6, pp. 832-840, 2003.
- [8] Koji Makita, Masayuki Kanbara, Naokazu Yokoya, "View management of annotations for wearable augmented reality," *Proceedings of the 2009 IEEE international conference on Multimedia and Expo (ICME)*, pp. 982-985, 2009.
- [9] 山下 淳, 葛岡 英明, 山崎 敬一, 山崎 晶子, 加藤 浩, 鈴木 栄幸, 三木 弘之, "相互モニタリング可能な遠隔会議システム," *日本バーチャルリアリティー学会論文誌*, vol. 4, no. 3, pp. 495-594, 1999.
- [10] 森川 治, 福井 幸男, 山下 樹里, 佐藤 滋, "人に優しい超鏡対話における指差し行為," *情報処理学会論文誌*, vol. 41, no. 5, pp. 1290-1297, 2000.
- [11] 早川 祐, 葛岡 英明, "遠隔会議における指差し支援の研究," *ヒューマンインタフェースシンポジウム論文集*, pp. 2322-2325, 2004.
- [12] Nobuchika Sakata, Takeshi Kurata, Takekazu Kato, Masakatsu Kourogi, Hideaki Kuzuoka, "WACL: Supporting telecommunications using wearable active camera with laser pointer," *Proceedings of 7th IEEE International Symposium on Wearable*

- Computers (ISWC)*, pp. 53-56, 2003.
- [13] Susanna Nilsson, Bjorn Johansson, Arne Jonsson, "Using AR to support cross-organisational collaboration in dynamic tasks," *Proceedings of 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2009)*, pp. 3-12, 2009.
 - [14] Harlan Hile, Radek Grzeszczuk, Alan Liu, Ramakrishna Vedantham, Jana Košecka, Gaetano Borriello, "Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning," *Proceedings of the 7th International Conference on Pervasive Computing*, pp. 59-76, 2009.
 - [15] Yoji Miyazaki, Toshiyuki Kamiya, "Pedestrian navigation system for mobile phones using panoramic landscape images," *International Symposium on Applications and the Internet (SAINT)*, pp. 23-27, 2006.
 - [16] Steffen Werner, Bernd Krieg-Brückner, Hanspeter A. Mallot, Karin Schweizer, Christian Freksa, "Spatial Cognition: The Role of Landmark, Route, and Survey Knowledge in Human and Robot Navigation," *Informatik aktuell, Springer*, pp. 41-50, 1997.
 - [17] Kazuyuki Fujita, Kazuki Takashima, Takayuki Tsukitani, Yuichi Itoh, Yoshifumi Kitamura, Fumio Kishino, "Anchored navigation: coupling panning operation with zooming and tilting based on the anchor point on a map," *Proceedings of Graphics Interface (GI)*, pp. 233-240, 2010.
 - [18] Antti Nurminen, Antti Oulasvirta, "Designing interactions for navigation in 3D mobile maps. In: Liqiu Meng, Alexander Zipf, and Stephan Winter (editors). Map-Based Mobile Services: Design, Interaction and Usability," *Springer. Lecture Notes in Geoinformation and Cartography*, pp. 198-227, 2008.
 - [19] Bjorn Schwerdtfeger, Rupert Reif, Willibald A. Gunthner, Gudrun Klinker, Daniel Hamacher, Lutz Schega, Irina Bockelmann, Fabian Doil, Johannes Tumler, "Pick-by-Vision: A first stress test," *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 115-124, 2009.
 - [20] Steven J. Henderson, Steven Feiner, "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret," *Proceedings of 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 135-144, 2009.
 - [21] J. Edward Swan, Joseph L. Gabbard, "Survey of User-Based Experimentation in Augmented Reality," *Proceedings of 1st International Conference on Virtual Reality, HCI International*, pp. 1-9, 2005.

- [22] Yoshinari Kameda, Taisuke Takemasa, Yuichi Ohta, "Outdoor see-through vision utilizing surveillance cameras," *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 151-160, 2004.
- [23] Peter Barnum, Yaser Sheikh, Ankur Datta, Takeo Kanade, "Dynamic Seethroughs: Synthesizing Hidden Views of Moving Objects," *Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 111-114, 2009.
- [24] Christian Sandor, Andrew Cunningham, Arindam Dey, Ville-Veikko Mattila, "An Augmented Reality X-Ray system based on visual saliency," *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 27-36, 2010.
- [25] 天目 隆平, 神原 誠之, 横矢 直和, "ウェアラブル拡張現実感システムのための注目オブジェクトへの直感的な注釈提示手法," *日本バーチャルリアリティ学会論文誌*, vol. 10, no. 3, pp. 305-312.
- [26] Christian Sandor, Arindam Dey, Andrew Cunningham, Sebastien Barbier, Ulrich Eck, Donald Urquhart, Michael R Marner, Graeme Jarvis, Sang Rhee, "Egocentric Space-Distorting Visualizations for Rapid Environment Exploration in Mobile Mixed Reality," *Proceedings of the 2010 IEEE Virtual Reality Conference (VR)*, pp. 47-50, 2010.
- [27] Richard Stoakley, Matthew J. Conway, "Virtual reality on a WIM: interactive worlds in miniature," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, pp. 265-272, 1995.
- [28] Chadwick A. Wingrave, Yonca Haciahetoglu, Doug A. Bowman, "Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM," *Proceedings of the IEEE conference on Virtual Reality (VR'06)*, pp. 11-16, 2006.
- [29] Tobias Höllerer, Steven Feiner, Drexel Hallaway, Blaine Bell, Marco Lanzagorta, Dennis Brown, Simon Julier, Yohan Baillot, Lawrence Rosenblum, "User interface management techniques for collaborative mobile augmented reality," *Computers & Graphics*, vol. 25, no. 5, pp. 799-810, 2001.
- [30] 大隈 隆史, 興梠 正克, 七田 洸一, 蔵田 武志, "科学ミュージアムガイドにおける三次元地図提示のための仮想視点制御と体験誘導コンテンツ提示の効果," *日本バーチャルリアリティ学会論文誌*, vol. 14, no. 2, 2009.
- [31] Antti Nurminen, "Mobile, hardware-accelerated urban 3D maps in 3G networks," *Proceedings of the 12th International Conference on 3D Web Technology (Web3D)*, pp. 7-16, 2007.
- [32] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Brian Curless, Steven M. Seitz, Richard Szeliski, "Building Rome in a Day," *Communications of the ACM*, vol. 54, no. 14,

- pp. 105-112, 2011.
- [33] Arnold Irscharaa, Christopher Zachb, Manfred Klopschitz, Horst Bischofa, "Large-scale, dense city reconstruction from user-contributed photos," *Computer Vision and Image Understanding*, vol. 116, no. 1, pp. 2-15, 2011.
 - [34] 清川 清, "複合現実型マルチユーザ・コラボレーション ～ヒューマンコミュニケーションの視点から～," *情報処理学会シンポジウム論文集*, pp. 49-58, 2002.
 - [35] Matthias Trapp, Lars Schneider, Christine Lehmann, Norman Holz, Jürgen Döllner, "Strategies for Visualizing 3D Points-of-Interest on Mobile Devices," *Journal of Location Based Services (JLBS)*, vol. 5, no. 2, pp. 79-99, 2011.
 - [36] Niklas Elmqvist, Philippas Tsigas, "A taxonomy of 3D occlusion management for visualization," *IEEE transactions on visualization and computer graphics*, vol. 14, no. 5, pp. 1095-1109, 2008.
 - [37] 村越 真, 若林 芳樹, "GIS と空間認知—進化する地図の科学—," 今古書院, 2008.
 - [38] Roger N. Shepard, Jacqueline Metzler, "Mental Rotation of Three-Dimensional Objects," *Science*, vol. 171, no. 3972, pp. 701-703, 1971.
 - [39] James C. Chung, "A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams," *Proceedings of Symposium on Interactive 3D Graphics*, pp. 193-196, 1992.
 - [40] David R. Koller, Mark R. Mine, Scott E. Hudson, "Head-Trackd Orbital Viewing: An Interaction Technique for Immersive Virtual Environments," *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pp. 81-82, 1996.
 - [41] Hirokazu Kato, Mark Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR)*, pp. 85-94, 1999.
 - [42] 山下 淳, 葛岡 英明, Paul Luff, 山崎 敬一, 晶子, "AgoraG: 遠隔操作型レーザポインタと書画カメラを備えた遠隔共同作業支援システム," *ヒューマン・インタフェース・シンポジウム (HIS)2001 論文集*, pp. 585-588, 2001.
 - [43] Yoshinari Kameda, Dao Thanh Ngoc, Itaru Kitahara, Yuichi Ohta, "MR based Visualization of Viewing Fields of Surveillance Cameras in Outdoor Scene," *The 1st International Workshop on Aware Computing(IWAC)*, pp. 733-736, 2009.
 - [44] 住谷 司, 亀田 能成, 大田 友一, "撮影空間における環境カメラ情報の複合現実感技術による可視化," *電子情報通信学会技術研究報告. HIP, ヒューマン情報処理*, vol. 109, no. 471, pp. 287-292, 2010.

- [45] Gerhard Reitmayr, Tom Drummond, "Going out: Robust Model-based Tracking for Outdoor Augmented Reality," *Proceedings of the 5th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 109-118, 2006.
- [46] Clemens Arth, Manfred Klopschitz, Gerhard Reitmayr, Dieter Schmalstieg, "Real-time self-localization from panoramic images on mobile devices," *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality*, pp. 37-46, 2011.
- [47] Georg Klein, David Murray, "Parallel Tracking and Mapping for Small AR Workspaces," *Proceedings of the 6th IEEE International Symposium on Mixed and Augmented Reality*, pp. 109-118, 2007.
- [48] Robert Castle, Georg Klein, Daid W. Murray, "Video-rate Localization in Multiple Maps for Wearable Augmented Reality," *Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers (ISWC)*, pp. 15-22, 2008.
- [49] 林 将之, 北原 格, 亀田 能成, 大田 友一, "画像から復元された3次元特徴点マップ群の世界座標系への統合," *電子情報通信学会総合大会講演論文集. 情報・システム(2)*, p. 188, 2010.
- [50] Hideyuki Tamura, Hirokazu Kato, "Proposal of international voluntary activities on establishing benchmark test schemes for AR/MR geometric registration and tracking methods," *Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 233-236, 2009.
- [51] Antoine Petit, Guillaume Caron, Hideaki Uchiyama, Eric March, "Evaluation of Model based Tracking with TrakMark Dataset," *2nd International Workshop on AR/MR Registration, Tracking and Benchmarking*, p. 5 pages, 2011.
- [52] Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, Yuichi Ohta, "A Study of Camera Tracking Evaluation on TrakMark Data-Set," *The 2nd International Workshop on AR/MR Registration, Tracking and Benchmarking (TrakMark)*, p. 5 pages, 2011.
- [53] Richard A. Newcombe, Andrew J. Davison, "Live dense reconstruction with a single moving camera," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1498-1505, 2010.
- [54] Christian Frueh, Siddharth Jain, Avidesh Zakhori, "Data Processing Algorithms for Generating Textured 3D Building Facade Meshes From Laser Scans and Camera Images," *Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission*, pp. 834-847, 2002.
- [55] "Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM," *Proceeding of IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 11-16, 2006.
- [56] David R. Koller, Mark R. Mine, Scott E. Hudson, "Head-tracked orbital viewing: an

- interaction technique for immersive virtual environments," *Proceedings of the 9th annual ACM symposium on User interface software and technology (UIST)*, pp. 81-82, 1996.
- [57] 高木 幸子, "コミュニケーションにおける表情および身体動作の役割," *早稲田大学大学院文学研究科紀要. 第1分冊, 哲学東洋哲学心理学社会学教育学*, vol. 51, pp. 25-36, 2006.
- [58] Christian Sandor, Andrew Cunningham, Ulrich Eck, Donald Urquhart, Graeme Jarvis, Arindam Dey, Sebastien Barbier, Michael R. Marner, Sang Rhee, "Egocentric Space-Distorting Visualizations for Rapid Environment Exploration in Mobile Mixed Reality," *Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR'09)*, pp. 211-212, 2009.

付録 A. PTAM の精度評価

PTAM を用いて屋外環境の特徴点マップを構築するにあたって、まず、屋外環境における PTAM による位置姿勢推定の精度を評価した。これには、TrakMark WG¹が公開している評価用データを用いた。TrakMark WG は、国際標準化に向けた AR/MR 向けのカメトラッキング手法のベンチマークテストの策定作業を行なっている [50]。これまでに、異なるシーンを対象とした 8 種類のパッケージが公開されている。各パッケージには、特定のシーンを撮影したいくつかの映像シーケンスと、撮影カメラの内部パラメータ、及び各フレームにおける外部パラメータの情報が含まれている。様々なカメラキャリブレーション方式をこのデータに対して実行することで、公平な比較が可能になると期待されている。

カメラキャリブレーション方式には、画像中の特徴点の 3 次元位置を求め、その再投影誤差を最小化するカメラ位置姿勢を推定する方式がしばしば用いられる。この方法では、最小化された再投影誤差がキャリブレーション精度を評価する指標である。しかし、この指標では、特徴点の 3 次元位置付近に仮想物体を置いた場合の投影誤差を知り予見することはできても、それ以外の位置に置いた場合に生じる投影誤差を予想することは難しい。

誤差評価の別の方式として、カメラの 3 次元位置と姿勢の誤差を距離差と角度差で表現する方法が提案されている [51] [52]。空間内でのカメラ位置姿勢が正確であれば、空間中のどの位置に仮想物体を置いた場合でも、正しく投影されることが保証される。しかし、推定したカメラ位置姿勢に誤差が含まれていた場合、その誤差が画面上での仮想物体の表示位置にどの程度の影響を与えるのかを予想するのは難しい。

そこで、画面上での表示位置の誤差を評価するために、シーン中の様々な位置に仮想的に 3 次元点群を配置し、その仮想点群を参照カメラ、推定カメラの各々の画像上に投影し、その投影点同士の画面上での距離を評価する方式が TrakMark WG の議論から生まれた。

空間内でのカメラ位置姿勢の誤差と、PTAM が利用した特徴点の数も同時に評価することで、3 次元点群の投影誤差による提案評価方式の意義について考察する。

¹ TrakMark –Benchmark Test Schemes for AR/MR Geometric Registration and Tracking Methods–, <http://ypcex.naist.jp/trakmark/>

3次元点群の投影誤差による評価

AR/MRにおける幾何学的整合性とは、仮想物体を現実映像中に正しい位置・姿勢で重畳表示することである。これは、現実空間における視点（カメラ）の位置姿勢を正確に推定することで達成できるため、カメラキャリブレーションの精度を評価する方法として、3次元空間内でのカメラ位置姿勢の精度を評価する手法がこれまで用いられてきた [51] [52]。

しかし、実際には、空間内でのカメラ位置姿勢よりも、実写映像中の仮想物体の表示位置がどの程度ずれるかが重要である。そこで、シーン中に仮想的に3次元点群を配置し、その点群を参照値と推定値のカメラ位置姿勢を用いて画像上に投影した場合の投影誤差を評価する方式を提案する。

この方式を、図 7-1 に示す透視投影モデルを用いて定式化する。前提として、カメラの位置姿勢は、世界座標系からカメラ座標系への 4×4 の変換行列 \mathbf{C} として表現し、参照値と推定値をそれぞれ \mathbf{C}_r , \mathbf{C}_e とする。カメラ行列 \mathbf{C}_r , \mathbf{C}_e で表現されるカメラをそれぞれ参照カメラ、推定カメラと呼ぶ。

まず図 1 上に示すように、参照カメラの正規化画像座標上に N 点の格子点 \mathbf{x}_i , $i = 1, \dots, N$ をとる。参照カメラの投影中心から、その格子点 \mathbf{x}_i を通る直線と、投影面に平行かつ投影中心から距離 d 離れた仮想平面 \mathbf{L} との交点 \mathbf{P}_i をシーン中の仮想的な3次元点群として配置する。これはつまり、 \mathbf{L} の近傍に仮想物体を重畳させることを期待している。この3次元点群 \mathbf{P}_i を推定カメラに投影し、正規化画像座標 \mathbf{y}_i を得る。

\mathbf{C}_r と \mathbf{C}_e が異なる場合、投影点 \mathbf{y}_i は格子点 \mathbf{x}_i からずれる。カメラの内部パラメータとレンズ歪みモデルを用いて \mathbf{x}_i , \mathbf{y}_i を画像座標に変換したものを \mathbf{x}'_i , \mathbf{y}'_i とする。このずれの平均を評価値とする。それらの距離 e_i の平均を平均投影誤差 E とする。

$$e_i = \|\mathbf{x}'_i - \mathbf{y}'_i\|$$

$$E = \frac{1}{N} \sum_i e_i$$

以上の手順によって、参照カメラから距離 d の仮想平面 \mathbf{L} 上においた仮想物体の表示位置のずれを求めることができる。参照カメラから仮想平面 \mathbf{L} までの距離 d を数パターン用意し、それぞれで平均投影誤差 E を求めることで、空間中の様々な奥行きに仮想物体を置いた時の表示位置ずれを評価することができる。

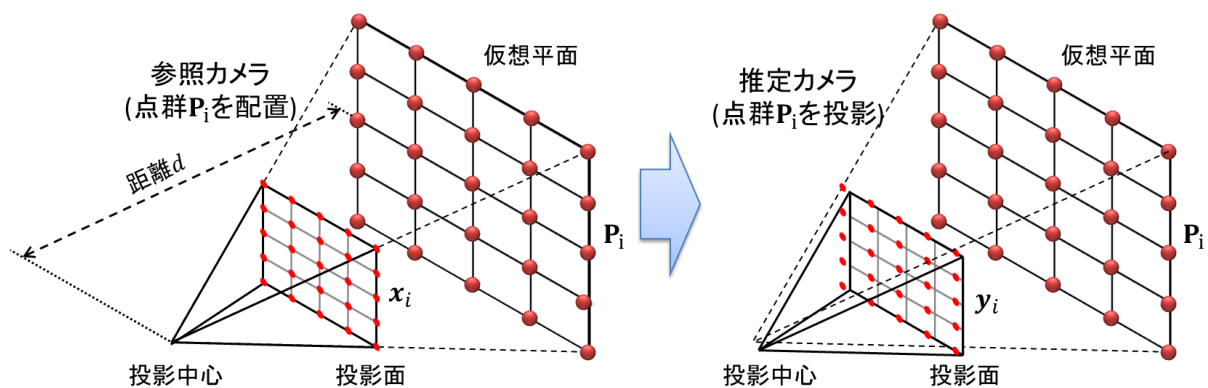


図 7-1 3次元点群の配置と、その投影誤差による位置姿勢の推定誤差評価

評価手順

以下の手順を用いて PTAM を提案方式で評価する。

1. 内部パラメータの設定

PTAM が用いる内部パラメータは、正規化焦点距離 (f_x, f_y), 正規化画像中心 (u_0, v_0) と Field of View モデル [53] の歪みパラメータ ω である。これらの情報は TrakMark データセットに含まれている。

2. ステレオ初期化に用いるフレームの設定

PTAM は初期の 3 次元特徴点マップを構築するためにステレオ視を用いる。そのために用いる 2 フレームを、シーケンス毎に決める必要がある。今回は先頭数十フレームの中から、初期化に成功しやすい 2 フレームを選択した。

3. PTAM の実行

映像シーケンスの各フレームにはタイムスタンプが付与されており、撮影時のフレームレートで入力することも可能であるが、本稿では PTAM の処理時間を確保するために、10 fps で入力を行った。

4. 世界座標系の変換

3.1.節の手順を用いて世界座標系の変換を行う。画像から推定されるカメラ位置はスケールが不定であるため、スケール付きユークリッド変換を行う。スケール s の計算方法は [52] と異なる。

5. 3次元点群の投影誤差による評価の実施

仮想平面を 4 パターン用意し、各フレームにおける平均投影誤差を求める。

世界座標系の変換

PTAM によって推定されるカメラ行列は、ステレオ初期化時に設定された世界座標系を基準としている。そのため、参照値が基準としている実空間に固定された世界座標系とは、原点、姿勢、及びスケールが異なる。そこで、スケール付きユークリッド変換を行うことで参照値と比較可能なカメラ行列を得る。

ステレオ初期化直後のカメラ位置姿勢は正しく推定されていると仮定する。その時のカメラ行列の推定値 \mathbf{C}_{e0} から参照値 \mathbf{C}_{t0} への変換を \mathbf{T} とする。

$$\mathbf{T} = \mathbf{C}_{t0}^{-1} \cdot s\mathbf{I} \cdot \mathbf{C}_{e0}$$

ここで、 \mathbf{I} は 4×4 の単位行列である。[52]ではカメラが初期位置とそこから最も離れた位置の距離からスケール s を求めている。しかし、カメラ位置の推定誤差による影響を受けやすいため、本稿では PTAM が仮定するステレオ視のベースライン長 d_s と、それに対応する参照値 d_r の比を用いる。

$$s = d_r/d_s$$

最終的に、第 i フレームにおけるカメラ行列 \mathbf{C}_{ei} から変換行列 \mathbf{T} を用いて、参照値 \mathbf{C}_{ti} と比較可能なカメラ行列 \mathbf{C}'_{ei} を得る。

$$\mathbf{C}'_{ei} = \mathbf{C}_{ei} \cdot \mathbf{T}^{-1}$$

評価結果

TrakMark WG が公開している、NAIST Campus Package 02 を用いて PTAM のキャリブレーション精度評価を行った結果を示す。このパッケージの映像シーケンスは、屋外歩行者が持つカメラから撮影された映像で、解像度は 1280x720 画素である。

今回はカメラの投影面上に 17x17 点の格子点を用いて配置した 3 次元点群を用いた ($N=289$)。

今回は実行結果として、空間内でのカメラ位置、姿勢（オイラー角）の参照値（TX, TY, TZ）と推定値（X, Y, Z）、3 次元点群の投影誤差に加えて、PTAM が構築した 3 次元マップの特徴点数 (nmp) とキーフレーム数 (nkf)、画像中の特徴点との 2D-3D 対応付けを試みた特徴点数 (apt)、実際に対応付けができた特徴点数 (fnd) を同時に観察して評価する。

Sequence 00 のサンプル画像（図 7-2）と結果（図 7-3, 図 7-4）を示す。このシーケンスでは、特徴点は十分な数が抽出された。シーケンス前半では再投影誤差が小さい特徴点が多数あり、そのような特徴点は主に遠方の建物上にあった。そのため、建物上に仮想物体を設置する場合は、画像内での幾何的整合性が確保されると考えられる。しかし、カメラ位置の推定誤差を見ると、姿勢や X 軸、Y 軸方向の位置の誤差はあまり大きくないものの、奥行き方向（Z 軸）の誤差が大

きくなっていることがわかる。

3次元点群の平均投影誤差は、撮影シーンの広さに合わせて $d = 2, 5, 10, 50$ [m]とした時の3次元点群を用いて評価した。その結果、投影誤差は非常に大きく、カメラ近くに用意した仮想点群ほど誤差が大きくなった。これは、ほとんどの特徴点が遠くの建物上にあったため、それより手前の空間で投影誤差が生じやすいためと考えられる。

これは、Z 軸方向のカメラ位置の推定誤差が大きく、その分だけ仮想点群が推定カメラから遠くなり、投影点が画面中央付近に集中する。どの距離の仮想点群においても、1800 フレーム付近までは、投影点が画面中央に集中した時の投影誤差に近づいている。その後、カメラの向きが変わり姿勢の推定誤差が大きくなったため、投影点が画面外に大きく外れ、投影誤差が増えている。

nkf や nmp を見ると、対応付けを試みた点数 apt は PTAM の参照実装の上限である 1000 点程あるにも関わらず、実際に対応付けができた数 fnd は少いことがわかる。



図 7-2 NAIST Campus Package 02 のサンプル画像

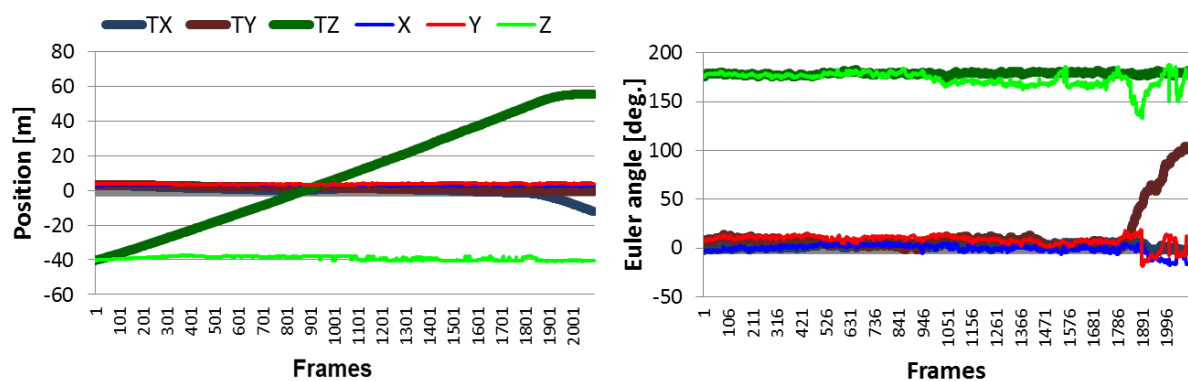


図 7-3 PTAM の位置姿勢推定の誤差

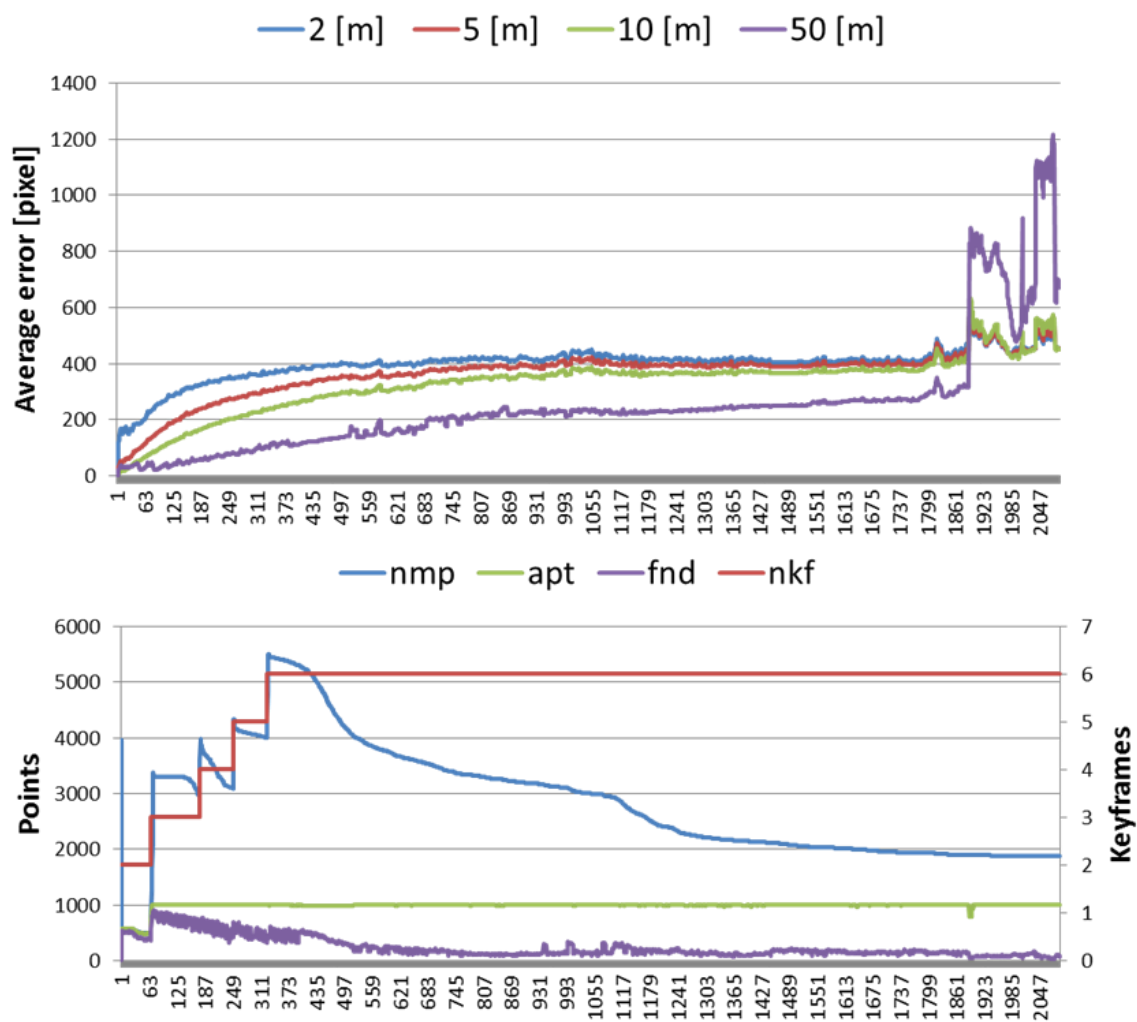


図 7-4 3次元点群の投影誤差

付録 B. 性能諸元

モバイルカメラ(Creative 製 VF0490)仕様

イメージセンサ	130 万画素カラーCMOS
対角視野角	85[deg] ±5%
焦点距離	パンフォーカス 15[cm]～∞
インタフェース	USB 2.0/1.1

慣性センサ(INTERSENSE 製 InertiaCube3)仕様

システム内容	3 軸方位角計測システム
自由度	3 軸(ヨー・ピッチ・ロール)
可動角度範囲	全方向 360[deg]
最大角速度	1200[deg/sec]
最小角速度	0[deg/sec]
RMS 精度	1[deg](ヨー), 0.25[deg](ピッチ・ロール)
RMS 角度分解能	0.03[deg]
データ更新レート	180[Hz]
データ遅延時間	2[ms](ホスト PC に依存)
シリアルレート	115.2[kbaud]
インタフェース	RS-232 シリアル
サイズ	26.2[mm]×39.2[mm]×14.8[mm]
重量	17.0[g]
電源	6[VDC], 40[mA]
動作可能環境温度	0～70[°C]

ユーザ端末(SONY 製 VGN-UX90PS)仕様

OS	Microsoft Windows XP 32bit
CPU	Intel Core Solo プロセッサ U1400(1.20 [GHz])
チップセット	945GMS Express チップセット
メインメモリ	512[MB](オンボード)
メモリバス	400[MHz]
ビデオメモリ	最大 128[MB](メインメモリと共有)
液晶ディスプレイ	4.5 型ワイド TFT カラー液晶(最大解像度 1024×600[pixel])
ワイヤレス通信	内臓(IEEE802.11a/b/g 準拠, WPA2 対応, Wi-Fi 適合) Bluetooth 機能内蔵(Bluetooth2.0+EDR 準拠)
消費電力	約 21[W]
外形寸法	150.2[mm]×32.2～38.2[mm]×95[mm](最大突起部 100.5[mm])
重量	約 520～532[g](付属バッテリーパック装着時)

ユーザ端末カメラ(SONY 製 VGN-UX90PS 背面内臓カメラ)仕様

システム内容	デジタルカメラ
イメージセンサ	1/4" SXGA プログレッシブ方式 CMOS センサ
レンズ	3 群 3 枚プラスチックレンズ(f=3.8[mm], F4)

ユーザ端末(Acer 製 ICONIA TAB W500S)仕様

OS	Microsoft Windows 7 32bit
CPU	AMD デュアルコア C-60 APU (1/1.33 [GHz])
チップセット	AMD A50M FCH
メインメモリ	2 [GB] DDR3
グラフィックスコア	AMD Radeon HD 6290
ビデオメモリ	最大 128[MB](メインメモリと共有)
液晶ディスプレイ	10.1 型ワイド(1280x800) 静電容量式マルチタッチパネル
ワイヤレス通信	無線 LAN (IEEE802.11b/g/n 準拠) Bluetooth (Bluetooth3.0+HS)
バッテリー駆動時間	約 6 時間
外形寸法	275 [mm]×190 [mm]×15.95 [mm]
重量	約 0.97[kg]

ユーザ端末カメラ(IDS 製 UI-1221LE-C-HQ)仕様

イメージセンサ	1/3" WVGA CMOS
有効画素数	752×480 [pixel]
レンズ	3 群 3 枚プラスチックレンズ(f=3.8[mm], F4)

公表論文リスト

査読付き雑誌論文

(筆頭)

1. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “An Installation of Privacy-Safe See-Through Vision,” *Procedia-Social and Behavioral Sciences*, Elsevier Ltd, Volume 2, Issue 1, pp.125-128, 2010.
2. 林将之, 北原格, 亀田能成, 大田友一, “広域屋外環境における複合現実型バーチャルジオラマインタフェースの評価”, 日本バーチャルリアリティ学会論文誌, Vol.16, No.2, pp.215-225, 2011.

(非筆頭)

3. 長嶋正典, 林将之, 北原格, 亀田能成, 大田友一, “録画された移動物体行動の複合現実型提示における一覧性の向上”, 日本バーチャルリアリティ学会論文誌, 日本バーチャルリアリティ学会論文誌, Vol.17, No.4, pp.447-456, 2012.

査読付き国際会議論文

(筆頭)

1. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “A User Study on Viewpoint Manipulation Methods for Diorama-Based Interface Utilizing Mobile Device Pose in Outdoor Environment,” 21st International Conference on Artificial Reality and Telexistence (ICAT2011), pp.41-46, 2011.

(非筆頭)

2. Yuichi Ohta, Yoshinari Kameda, Itaru Kitahara, Masayuki Hayashi, and Shinya Yamazaki, “See-Through Vision: A Visual Augmentation Method for Sensing-Web,” International Conference on Information Processing and Management of Uncertainty in Knowledge-Based System (IPMU2010), 11 pages, 2010.

その他の公表論文

1. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “Pointing and Sharing 3D Space By Using a Virtual Diorama for Outdoor Mixed Reality,” The 1st Korea-Japan Workshop on Mixed Reality, 6 pages, 2008.
2. 林将之, 北原格, 亀田能成, 大田友一, “仮想俯瞰模型を用いた3次元空間の指示・共有システム”, 日本バーチャルリアリティ学会第13回大会論文集, pp.644-647, 2008.
3. 林将之, 北原格, 亀田能成, 大田友一, “モバイル端末画像を用いたクライアント・サーバ

型の PTAM 処理”，電子情報通信学会 2009 年総合大会, 1 page, 2009.

4. 林将之, 北原格, 亀田能成, 大田友一, “仮想俯瞰模型を用いた 3 次元位置共有方式における情報提示手法の検討と評価”, 日本バーチャルリアリティ学会第 14 回大会, 4 pages, 2009.
5. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “3D Pointing Interface by using Virtual Diorama for Attention Sharing,” “Let’s Go Out” Workshop in conjunction with ISMAR2009, 2 pages, 2009.
6. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “An Installation of Privacy-Safe See-Through Vision,” International Conference on Security Camera Network, Privacy Protection and Community Safety, 2 pages, 2009.
7. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “Maneuvering of Virtual Diorama for Finding Attention in Wide Area Outdoor Workspace,” The 3rd Korea-Japan Workshop on Mixed Reality, (slides), 2010.
8. 林将之, 北原格, 亀田能成, 大田友一, “画像から復元された 3 次元特徴点マップ群の世界座標系への統合”, 電子情報通信学会 2010 年総合大会, 1 page, 2010.
9. 林将之, 北原格, 亀田能成, 大田友一, “AR/MR デモンストラショの再現性を保証するソフトウェア環境の構築—Casper Cartridge プロジェクター”, 電子情報通信学会 技術研究報告 MVE, Vol.111, No.235, pp.103-108, 2011.
10. Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, “A Study of Camera Tracking Evaluation on TrakMark Data-Set,” The 2nd International Workshop on AR/MR Registration, Tracking and Benchmarking (TrakMark2011), 5 pages, 2011.
11. 林将之, 北原格, 亀田能成, 大田友一, “屋外におけるジオラマ型インタフェースを用いたパン・チルトカメラの可視化”, 電子情報通信学会 2012 年総合大会, 1page, 2012.
12. 林将之, 北原格, 亀田能成, 大田友一, “シーン中の 3 次元点群の投影誤差による AR/MR 向けカメラキャリブレーションの精度評価”, 電子情報通信学会 技術研究報告 MVE, Vol.112, No.221, pp.43-48, 2012.
13. 林将之, 北原格, 亀田能成, 大田友一, “屋外カメラ設置におけるポーズ決定タスクによる複合現実型ジオラマインタフェースの評価”, 電子情報通信学会 技術研究報告 MVE, Vol.113, No.51, pp.45-50, 2013.

受賞

1. 2011 年 9 月 第 15 回 パターン認識・メディア理解研究会 アルゴリズムコンテスト 入賞, 鳥屋剛毅, 出嶋康之, 山桐靖史, 林将之, 小林直樹, “極座標画像による指先抽出と凸包処理を併用したじゃんけんの勝敗判定”, 電子情報通信学会 パターン認識・メディア理解研究専門委員会.

2. 2012 年 9 月 日本バーチャルリアリティ学会複合現実型研究委員会 平成 23 年度 SIG-MR 賞, 林将之, 北原格, 亀田能成, 大田友一, “AR/MR デモンストレーションの再現性を保証するソフトウェア環境の構築—Casper Cartridge プロジェクター”, 電子情報通信学会 技術研究報告 MVE.

研究助成金

1. 平成 23 年度～平成 24 年度 科学研究費補助金 特別研究員奨励費 (課題番号 23・310)