

NAIST-IS-DD0561019

博士論文

視点位置計測に基づく運動視差3次元ディスプレイの
開発と立体感の評価

末永 剛

2008年2月7日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

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

未永 剛

審査委員：

小笠原 司 教授	(主指導教員)
松本 健一 教授	(副指導教員)
松本 吉央 客員准教授	(副指導教員)
横矢 直和 教授	(委員)

視点位置計測に基づく運動視差3次元ディスプレイの 開発と立体感の評価*

末永 剛

内容梗概

3次元ディスプレイはより高い臨場感や場の共有感などが得られると期待されており、ロボットの操作や医療、アミューズメントなど様々な応用分野において利用が検討されている。しかし日常で利用されているテレビやコンピュータのモニタについては未だ2次的にしか利用されておらず、いくつか製品化されている裸眼立体視ディスプレイも高価なため手軽に利用できるものではない。こういった日常環境においても利用できる3次元表示システムが実現できれば、より手軽に3次元情報を利用することが可能になり、新たな情報提示や情報共有の可能性が広がると考えられる。

本論文では人間の奥行き知覚要因の一つである運動視差を用いた立体提示に焦点を当てる。そして非接触な顔情報計測に基づくモデルベースおよびイメージベースにおける運動視差3次元ディスプレイを提案し、コンピュータグラフィックスで作成された物体(以下、3D CG)や実世界で撮影された画像からの立体提示を行う。またこういった分野でたびたび無視される奥行き知覚実験を通してシステムを評価する。

はじめに、運動視差提示3次元ディスプレイを構築する上で必要となるアルゴリズムについて議論する。ヘッドマウントディスプレイのような人間の動きに追従するモニタ面に提示する場合と通常モニタのように環境に固定されたモニタ面に提示する場合では、視点位置とモニタの位置関係に違いが生じるため、知覚

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 博士論文, NAIST-IS-DD0561019, 2008年2月7日.

される画像の歪みを想定しなければならない。そこでモデルベースシステムでは3次元モデルを、イメージベースシステムでは実写画像を射影変換することにより任意視点に対応した画像を提示する。

次に、モデルベースやイメージベースにおける運動視差3次元ディスプレイの実装について述べる。ユーザの視点位置はモニタ上部に配置されたステレオカメラから計測され視点位置に対応した画像を生成しユーザに提示する。イメージベースにおいては、6軸マニピュレータを用いて複数視点画像の撮影を行い、事前に収集した画像群を用いた例とオンラインで動作させた例において実装した。

最後に、奥行き知覚実験を通して実装したシステムの評価を行う。モデルベースにおいては絶対距離知覚実験を行い、過去の研究の通り、より遠くの物体ほど知覚される距離が過小評価されるという知見が得られたものの、物体の前後関係や距離の違いを認識できることが確認され簡易なシステムとしては十分使用可能であることが示された。イメージベースにおいては、離散的に取得される画像の視野分解能についてや、アクティブカメラにおいて想定される提示遅れの許容についての検討を行い、通常の平面モニタを利用して非接触な視点位置計測に基づいた立体提示を行うことの実現可能性を示した。

この結果は通常の平面モニタを利用し、非接触な視点位置計測に基づいた立体提示を行うことの実現性を示すもので、今まで主に2次元のかまたはマウス操作により3次元データを扱っていたウェブブラウザ、3D CADなどのソフトウェア、さらにはテレビなどにおいて立体情報を扱った新しいアプリケーション制作の指針となるものである。

キーワード

3次元ディスプレイ, 運動視差, 非接触, 顔情報計測, 立体感評価

Development and Evaluation of 3D Display Based on Motion Parallax Using Viewpoint Measurement*

Tsuyoshi Suenaga

Abstract

It is expected for 3D displays to induce higher realistic sensation. They are adopted in various applications such as tele-operation, tele-surgery, 3D designing and amusement. Some glass-less stereoscopic displays based on lenticular lens or parallax barriers are commercially available, however they are still expensive to be a popular device.

In this paper, A 3D display based on motion parallax using non-contact face tracking system is proposed. A 3D model-based and an image-based approach is implemented. Through depth perception experiments, which are a topic often neglected, the feasibility of the system is confirmed. First, the algorithm of a 3D display based on motion parallax is proposed. It is necessary to warp the images according to the position of the viewpoint when the standard monitor is used compared with the head mounted display. The image that corresponds to the viewpoint of the user is generated from the projection of the 3D model or one of the multi-view images.

Next, a 3D model-based and an image-based 3D display is implemented and discussed. The 3D viewpoint of the user is measured using a real-time non-contact face tracking system using a stereo camera pair at the top of the monitor. In

*Doctoral Dissertation, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0561019, February 7, 2008.

the image-based approach, an off-line and an on-line system is implemented. In the off-line system, the multi-view images are obtained in advance by a camera mounted on a manipulator. It can provide the 3D information without a delay. In the on-line system, a camera mounted on a manipulator are controlled according to the position of the viewpoint.

Lastly, through experiments, the feasibility of the system was confirmed. In the 3D model-based approach, the error grew as the depth became larger. However, it was confirmed that the system was able to induce depth perception on the subject. In the off-line image-based system, the necessary number of the presented images were investigated. Moreover, the maximum delay allowed for presenting images were investigated to realize an on-line version of the system.

Keywords:

3D display, motion parallax, non-contact, face tracking, depth perception

目次

1. 序論	1
1.1 研究の背景	1
1.2 立体感と人間の奥行き知覚手掛かり	2
1.3 3次元ディスプレイの現状	4
1.4 研究の目的と意義	7
1.5 論文の構成	7
2. 運動視差による立体知覚	9
2.1 本章の趣旨	9
2.2 運動視差提示手法の分類	9
2.2.1 運動視差3次元ディスプレイの視点位置計測	9
2.2.2 3次元映像の提示方法の分類	10
2.3 運動視差3次元ディスプレイの評価	12
2.4 本研究の位置付け	12
2.5 本章の結論	13
3. 非接触な顔情報計測に基づく運動視差3次元ディスプレイの提案	14
3.1 本章の趣旨	14
3.2 顔情報計測に基づく視点位置計測	14
3.3 運動視差提示のための透視投影変換	15
3.3.1 視点位置により知覚される映像の違い	15
3.3.2 固定モニタへの射影変換	17
3.4 本章の結論	22
4. 非接触な顔情報計測に基づく運動視差3次元ディスプレイの実装	23
4.1 本章の趣旨	23
4.2 実装1：モデルベースによる運動視差提示3次元ディスプレイ	23
4.2.1 モデルベース3次元ディスプレイのシステム構成	23
4.2.2 回転ステージによる計測範囲拡大	24

4.3	実装2：多視点画像データベースを用いたイメージベース運動視差提示3次元ディスプレイ	28
4.3.1	マニピュレータによる多視点実画像の撮像	28
4.3.2	多視点実画像からの運動視差提示	29
4.4	実装3：オンラインでのイメージベース運動視差提示3次元ディスプレイ	32
4.5	イメージベースによる運動視差3次元ディスプレイの提示画像評価	38
4.6	本章の結論	40
5.	奥行き知覚に基づく評価実験	41
5.1	本章の趣旨	41
5.2	モデルベースによる運動視差3次元ディスプレイの評価	41
5.2.1	運動視差のみによる絶対距離知覚	41
5.2.2	絵画的要因との統合による絶対距離知覚	45
5.3	イメージベースシステムによる物体配置の知覚精度評価実験	50
5.3.1	多視点実画像の蓄積による運動視差提示実験方法	50
5.3.2	画像の提示間隔の評価	55
5.3.3	提示遅れによる知覚精度の評価	59
5.4	本章の結論	61
6.	結論	62
6.1	まとめ	62
6.2	今後の課題と展望	63
	謝辞	65
	参考文献	68
	付録	76
A.	顔情報計測システム	76

B. 7 軸マニピュレータ PA10 の仕様	82
C. 6 軸マニピュレータの逆運動学	85
D. 回転行列から回転要素の導出	86
D.1 ロールピッチヨー角への変換式	86
D.2 オイラー角への変換式	88

目 次

1	奥行き知覚手掛かり	3
2	観察者からの距離における奥行き知覚度合い	6
3	顔情報計測結果	15
4	実世界でモニタ枠越しに景色を見た時の見えの変化	16
5	非対称な透視変換	19
6	視点に対応した画像取得と射影変換	21
7	モデルベース3次元ディスプレイのハードウェア構成図	24
8	モデルベース3次元ディスプレイの運動視差提示の様子	26
9	モデルベース3次元ディスプレイの回転ステージを用いない運動視差提示結果	27
10	モデルベース3次元ディスプレイの回転ステージを用いた運動視差提示結果	27
11	画像データベースを用いた運動視差提示3次元ディスプレイの概要	28
12	多視点画像データベースの構造	30
13	多視点画像データベースを用いた運動視差提示の様子	31
14	多視点画像データベースを用いた運動視差提示の結果	31
15	オンラインによる運動視差提示3次元ディスプレイの処理の流れ	34
16	オンラインによる運動視差提示の様子	35
17	オンラインによる運動視差提示結果(4秒間隔)	36
18	マニピュレータ左右移動の遅れ	37
19	マニピュレータ上下移動の遅れ	37
20	マニピュレータ前後移動の遅れ	37
21	トラッキング用マーカを取り付けたマニピュレータ	38
22	提示画像の整合性の確認実験結果 (a) キャプチャ画像, (b) 生成画像 (射影画像), (c) ユーザの視点位置から観測される画像	39
23	絶対距離知覚実験での描画画像	43
24	実験で使用した視覚対象	43
25	片眼両眼における知覚距離の平均	45

26	壁の描画と合わせた絶対距離知覚実験での描画画像	46
27	壁ありでの知覚距離の平均	47
28	イメージベースシステムによる物体配置の知覚評価実験システム	52
29	マーカ付き実験用オブジェクトと配置物体	53
30	実験用オブジェクトの位置誤差計測インタフェース	53
31	実験で使用した画像の例	54
32	画像提示間隔の物体配置評価	56
33	画像提示間隔のアンケートによる主観評価	57
34	画像提示間隔に対する物体知覚範囲	58
35	画像提示遅延の物体配置評価	60
36	画像提示遅延のアンケートによる主観評価	60
37	頭部位置視線計測システムのソフトウェア構成図	77
38	登録した顔画像	78
39	特徴領域のテンプレート画像	78
40	顔全体のテンプレート画像	79
41	視線モデル	81
42	顔，視線計測結果	81
43	マニピュレータ PA10 の関節構造	83
44	マニピュレータ PA10 の回転軸	84
45	第 3 関節を固定したマニピュレータ PA10 の関節構造	85

表 目 次

1	多視点画像データベースを用いたシステムの処理時間，画像サイズ 640 × 480	29
2	処理時間	33
3	片眼両眼における知覚距離の正規化誤差	45
4	壁あり・なしでの知覚距離の正規化誤差	46
5	両眼利用時の正規化誤差 (各被験者)	49

6	2 要因分散分析結果 (各被験者)	49
7	立体感のライアン法による多重比較の結果	57
8	配置しやすさのライアン法による多重比較の結果	57
9	立体感のライアン法による多重比較の結果	60
10	マニピュレータ PA10 の仕様	82

1. 序論

1.1 研究の背景

ここ十数年の間にユーザインタフェースは大きく進化しており、マウス操作などによる Graphical User Interface などから、より直感的な操作である音声やジェスチャーといった言語・非言語情報による Perceptual User Interface へと移行してきている。それにともない、ディスプレイデバイスも直感的な利用を目指し、物や景色などとも立体的なものをそのまま立体的に表示させる 3 次元ディスプレイの研究、開発が盛んである。3 次元ディスプレイは、CAD や 3D CG のモデリングなどにおいて操作性の向上が見込まれたり、遠隔コミュニケーションや遠隔操作においてもより高い臨場感が得られると期待されており、ロボットの操作や医療、アミューズメントなど様々な応用分野において利用が検討されている。

3 次元のコンテンツにおいては、貴重な文化財をデジタル化し、蓄積、整備することでデータベースとして保存するデジタルアーカイブが盛んに行われており、世界各地の遺跡や仏像など様々な物が 3 次元アーカイブ化されている [1]。さらに、それらアーカイブ化された物をネットワークなどを通じて閲覧可能にするデジタルミュージアム、ユビキタスミュージアム [2] などの取り組みも実施されている。また身近なところではインターネットを通じた車やパソコンなどの新製品の紹介においてマウス操作により製品を立体的に閲覧可能にしているものも多く見られ、今後 3 次元情報を利用する機会は増えていくことが想定される。ネットワークによるコミュニケーションでは、ビデオ会議システムやビデオチャットなどを手軽に利用できるようになってきている。さらにテムザック社の家庭向け留守番ロボットにはカメラが搭載され、外出先から自宅の様子を携帯電話の画面越しに確認できるシステムなども登場している。将来的には相手の顔や遠隔地の情景が見えるだけでなく、映像が立体表示され、遠くにいながらも同じ空間で会話や作業をしているような感覚を与えることも期待される。

このように 3 次元提示のためのディスプレイデバイスの開発やコンテンツが増えていく中、日常的に広く利用されているテレビやコンピュータのモニタについては 2 次元的にしか利用されていないのが現状である。こういった日常環境にお

いても利用できる3次元表示システムが実現できれば、より手軽に3次元情報を利用することが可能になると考えられる。

1.2 立体感と人間の奥行き知覚手掛かり

立体感とは、平面的ではなく、奥行き・厚みなどのある感じのことを指し、人が立体に感じる奥行き手がかりには、運動視差や両眼視差、輻輳、焦点の遠近調節、視野内の高さ、大気遠近法、重なり合い、陰影、相対的な大きさなど様々なものがある[3][4][5][6]。図1に運動視差、両眼視差、輻輳、焦点の遠近調節の例を示す。図1(a)の運動視差は頭を動かすことで得られる像の変化である。電車の車窓から見える近くの景色が速く動き、遠くの景色がゆっくり動くのもこの運動視差によるものであるが、厳密には人間が自分で頭を動かすことによる頭部動きに追従した像の変化と頭部は動かず周りの景色が動くこのによる像の変化は区別される。図1(b)の両眼視差と図1(c)の輻輳はどちらも両眼要因などと分類されることもあり非常に類似しているが、両眼視差が左右の目で知覚した像の違いであるのに対して、輻輳は眼球を内側に向けることで発生する眼筋の変化である。ステレオグラムはこの両眼視差を利用した擬似立体視と言える。図1(d)の焦点の遠近調節とは眼球上のレンズの緊張と弛緩であり、カメラのピント調節に相当する。これらの運動視差や両眼視差、輻輳、焦点の遠近調整は人間の生理的な要因であり、人間の感覚器官に密接な関係がある。これに対してオクルージョンや視野内の高さ、相対的な大きさ、空気遠近感などは静止画に対しても効果があるため絵画的要因などと呼ばれることもある。3次元ディスプレイはこれらの奥行き手がかりを観察者に提示することで実現されるが、特に生理的な要因である運動視差、両眼視差、輻輳、焦点の遠近調節をうまく観察者に提示する必要がある。全ての奥行き知覚手掛かりを再現できれば一切違和感を感じさせずに立体感を提示することが可能になるが、現状でそれは困難であり一部を提示することで擬似的に立体感を提示することになる。

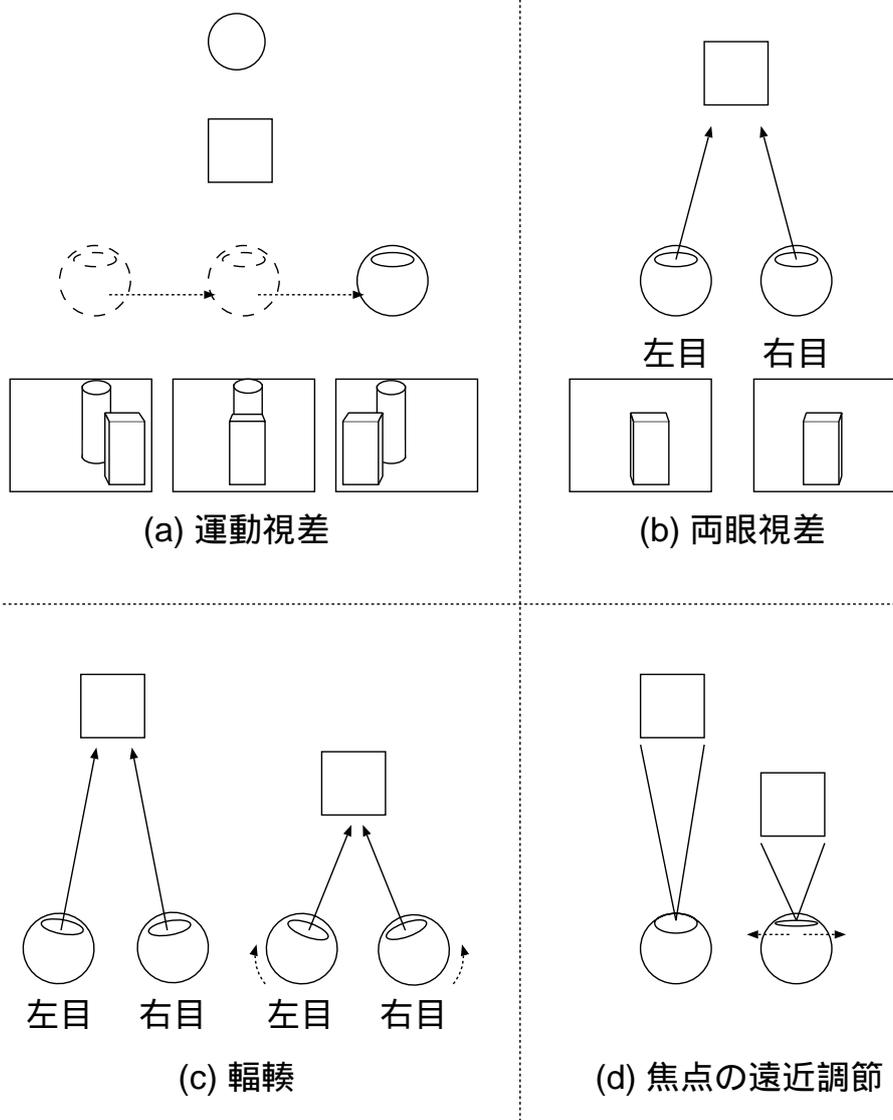


図 1 奥行き知覚手掛かり

1.3 3次元ディスプレイの現状

ユーザに対して立体提示を行う上で比較的簡単な実現方法としては赤青眼鏡を利用するアナグリフ方式やシャッタ眼鏡方式，偏光眼鏡方式などが挙げられる．アナグリフ方式は二つの異なった色のカラーフィルタを用い，フィルタの色以外の色を除くことで両目に別々の像を提示する．右目用，左目用の画像の配色調整のみで作成でき印刷物などにも利用できるため，本やポスターなど様々な場面で見掛けることも多い．ただし，色の違いで立体感を与えるため，使用できる色に制限が生じてしまう．シャッタ眼鏡はディスプレイのリフレッシュレートと同期させ左右のシャッタを交互に閉じることで左右の目に別々の像を提示するものであり，フレームレートが半分となる．偏光眼鏡は二つの重なった別々の画像に偏光フィルタをかけ，右目と左目で別の向きに取り付けた偏光フィルムを通して映像を見ることで左右別々の像を知覚することができる．IMAXシアターなどの3D映画館では初期の頃はシャッタ眼鏡が使用され，最近では偏光眼鏡が多く用いられている．近年は10万円前後の偏光式2D/3D液晶モニタも市販されている．これら特殊な眼鏡を着用した場合は両眼視差の提示により比較的簡単に立体知覚が行われるが，眼鏡の着用を必要とするため，ユーザへの拘束感は否めない．

またヘッドマウントディスプレイ(以下，HMD)を用いた場合には直接，ディスプレイデバイスを装着しており，ユーザの顔向き移動によらず映像を提示することができる．またジャイロセンサなどを取り付け，頭部位置計測を行うことで見る向きの制限なく自由な方向の映像を提示することが可能となり，ウォークスルーなどによく用いられる．最近では小型化，軽量化，高解像度化が進んでいるが，眼鏡方式と同様に装着の必要がある．

装着を必要としない両眼視差に基づいた3次元ディスプレイとしては，シャープやNEC，三洋電機，SeeReal，日商エレクトロニクス(SynthaGram)などから裸眼立体視ディスプレイが販売されている．これらはレンチキュラーレンズや偏光フィルム，微小なスリットなどを利用した，視差バリア方式[7]やインテグラルイメージング方式などが存在する．視差バリア方式には見る位置が固定されるといった問題が存在するが，Perlinらは，可動視差バリアを計測されたユーザの頭部位置に応じて制御することで，この問題を解決している[8]．インテグラルイ

メーキング方式では，ディスプレイにより違った映像を見ることができ視差数が存在し，比較的自由に頭を動かすことが可能である．しかし，視差数を増やすと解像度が低くなるという欠点が存在する上，高価なため手軽に利用できるものではない．

また規模の大きいものとしては，観察者の正面を大きなスクリーンで囲い没入させる視覚システムがあり，Deep Vision Display Wall[9] や Cylindra, D-vision[10] などが挙げられる．また柴野らは，CyberDome[11] という半球面のドーム型の没入型視覚システムを構築している．CyberDome は歪補正をソフトウェアで実現することで，特殊なレンズを必要とせず，投影手段配置をパラメータ化することにより，小型から大型まで任意サイズのドームスクリーンの実現を可能としている．部屋全体に対して映像を投影し，ユーザがその中に入るタイプの没入型ディスプレイもある．これには，CAVE[12]，CABIN[13]，COSMOS[14] などがあり，キューブ状の部屋に対して，それぞれ，4面，5面，6面に映像を投影可能なものである．これら大型の視覚システムは，両眼立体視用の眼鏡をかけたり，頭部位置計測用のセンサを頭や眼鏡に付けることで位置計測を行い，三次元ディスプレイとして利用されている．同様にユーザをディスプレイが囲うシステムの一つに田中らの Twister[15] がある．Twister は円筒状をしており，まわりを囲った円筒のディスプレイが回転しながら映像を表示させ，可動視差バリアにより裸眼立体提示を行っている．これらのシステムは大規模であり，美術館や博物館など公共の施設に導入されることはあるが一般家庭で手軽に利用するのは困難である．

これまでのシステムは使用人数が一人に限られていたが，複数人でも利用可能にした三次元ディスプレイも存在している．Agrawala らはシャッター眼鏡を用いた時分割方式による多人数化を提案しており [16]，また北村らは 4 人まで同時利用が可能な IllusionHole を構築している [17]．日商エレクトロニクス株式会社からは半球状ポリウムディスプレイ Perspecta(Actuality Systems, Inc.) が販売されている．また島田らはレーザをプラズマ化して実際の 3 次元空間中に点を発光させることで映像を表示する 3 次元映像装置を開発している [18]．この方法では両眼視差だけでなく輻輳や運動視差，焦点の遠近調節の提示を完璧に実現することが可能であり，どこからでも何人でも見ることが可能である．しかし，このデ

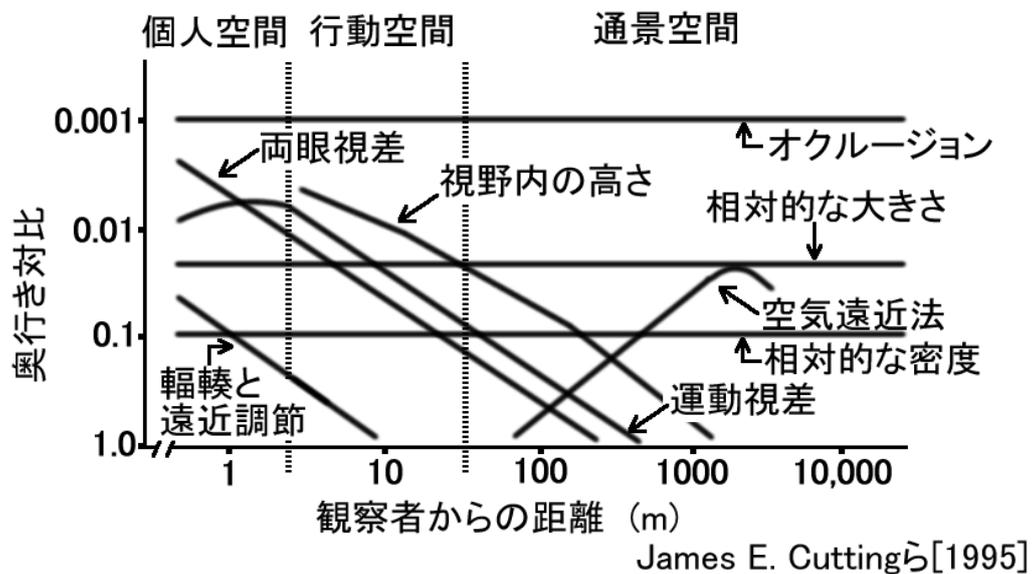


図 2 観察者からの距離における奥行き知覚度合い

バイスも特殊なモニタと同様に解像度の低さに問題があり、また実際にレーザが出るため安全対策などを必要とする。これら以外にも日立製作所の高速回転するスクリーンを利用したボリューム3次元ディスプレイ Transpost や12枚の固定された鏡を利用した持ち運び可能なボリューム3次元ディスプレイ、SeeReal社のホログラム方式ディスプレイなど様々あり、Doug A. Bowmanら[6]やL. Onuralら[19]は通常モニタやHMD、3次元ディスプレイに関わる技術をディスプレイデバイスの種類について分類している。

これらのように3次元ディスプレイは様々な視点から多くの研究が生まれているが、規模の大きさやコストにまだ課題が多い。また、液晶モニタやプラズマディスプレイなど2次元モニタが長い間広く普及しており、既にある2次元コンテンツとの互換性についても課題である。

1.4 研究の目的と意義

James E. Cutting らは様々な立体知覚要因における観察距離に応じた奥行き
の知覚度合(図2)を報告しており[20],これによれば,観察者からの距離が近い場
合,運動視差は両眼視差とほぼ同等の影響があり,輻輳や遠近調節よりも影響が
大きいことが分かる.よって運動視差を利用した3次元ディスプレイも簡易なシ
ステムとしては有効な手段であると考えられる.運動視差を単独で利用した3次
元表示システムは,観察者の視点位置を計測して視点位置に対応した画像を提示
することで実現が可能であり,ディスプレイデバイスとしては特別なものを必要
としない.そのため,今まで利用されている平面ディスプレイをそのまま利用す
ることが可能であり,従来の2次元コンテンツとの親和性が高い.また視点位置
計測に小型化低価格化が進んでいるカメラを利用することでコスト面においても
メリットであり,特別なデバイスを人体に装着しなくても利用可能である.カメ
ラを用いて視点位置を計測するためには,ユーザの顔を計測する必要があるが,
計測技術に関してもデジタルカメラに顔認識が搭載されるなど比較的,一般化し
つつあり今後も低価格化が期待される.

そこで本研究では,非接触な顔情報計測に基づく運動視差提示による3次元ディ
スプレイの提案を行う.まず3D CG モデルを利用したモデルベースシステムの
構築を行い,平面モニタに対しての射影変換の確認を行う.次に視点移動に追従
したカメラからの画像を利用したイメージベースシステムの構築を行い,実写画
像による3次元提示を行う.また従来の3次元ディスプレイでは,システムを利
用したときの奥行き感の評価についてはたびたび無視されてきた.本システムにお
いては,ユーザへの奥行き知覚実験などを通して実装したシステムの有効性の確
認を行い,通常平面モニタを利用して非装着な視点位置計測に基づいた立体提
示を行うことの実現可能性を示す.

1.5 論文の構成

本論文は以下の6章から構成されている.

第1章 : 研究背景として3次元ディスプレイの現状と有用性を述べ, 3次元ディスプレイの現状の問題点を明確にし, 本研究の目的を述べた.

第2章 : 運動視差に基づく立体提示を行う上での必要な技術と関連研究について紹介し, それぞれの研究との比較から本研究の位置づけを明確にする.

第3章 : 非接触な顔情報計測に基づく運動視差提示3次元ディスプレイの提案を行い, 実装する上で必要となるアルゴリズムについて議論する. まずベースとなる視点計測システムについて説明し, その後, 固定モニタに提示する上で発生する歪みについて説明し, 補正のための射影変換について説明する.

第4章 : モデルベースやイメージベースにおける運動視差3次元ディスプレイの実装について述べる. ユーザの視点位置はステレオカメラにより計測され視点位置に対応した画像を生成しユーザに提示する. モデルベースの場合は読み込んだ3D CGモデルを, イメージベースの場合にはある視点の画像をそれぞれ射影変換することでユーザ位置に対して整合性のとれた画像を生成する. 多視点画像の撮影は, 6軸マニピュレータを用いて行い, 事前に収集した多視点画像データベースを用いた例とオンラインで動作させた例について説明する.

第5章 : 奥行き知覚実験を通して実装したシステムの評価を行う. モデルベースにおいては絶対距離知覚実験を行い, 3次元ディスプレイを通して知覚された物体についてユーザが感じる距離について検証を行った. イメージベースにおいては, 多視点画像データベース方式において離散的に取得される画像の提示間隔についてや, オンラインシステムにおいて想定される提示遅れの許容についての検討を行った.

第6章 : 本論文を結ぶ. 本研究を通して得られた成果を総括し, 今後の展開について述べる.

2. 運動視差による立体知覚

2.1 本章の趣旨

本章では運動視差に基づく立体提示を行う上での必要な技術と技術毎における関連研究を述べ、それぞれの研究との比較から本研究の位置づけを明確にする。

2.2 運動視差提示手法の分類

ユーザに対して運動視差を提示する場合、視点位置に応じた映像をモニタに表示することになる。そのため、ユーザの視点である目の3次元位置を計測すること、また視点に対応した画像を生成すること、この2つを行う必要がある。以下、これらの研究における代表的な研究について述べる。

2.2.1 運動視差3次元ディスプレイの視点位置計測

位置を計測する研究は古くから行われており、磁気センサ、レーザセンサ、赤外線センサ、カメラなど様々な位置計測技術が開発されている。その中でも3次元ディスプレイにおいてよく利用されるものとしては磁気センサやカメラが挙げられる。

磁気センサの代表的な物としては Polhemus センサがあり、磁気発生装置と磁気センサがセットになり装置間の相対位置・姿勢を検出することが可能である。比較的安定して位置を取得することが出来るが、装着の必要があり、また磁場の歪みを発生させるため金属の近くでは使用できない。磁気センサは主に、HMDでの映像提示や、CAVE[12]などの没入ディスプレイにおいて多く用いられている。館らは、オブジェクト指向型ディスプレイとして *MEDIA*³(Media Cube)[21]を提案している。このシステムも磁気センサを利用して頭部位置の計測を行っており、小型液晶モニタを外向きに4枚組み合わせた直方体状のディスプレイであり、モニタに囲まれた箱の中の3D CGを立体的に見ることが可能である。

次にカメラを用いた研究として、Mulderらは頭部に取り付けたマーカをステレオカメラでトラッキングし、ハーフミラーを用いたAR環境での3次元ディスプ

レイを構築している [22] . また , Pastoor らはカメラを用いた視線入力可能な 3 次元ディスプレイを構築し , それを使ったコンピュータインタフェースを提案している [23] . このシステムは単眼カラーカメラによる頭部位置計測を行っている . 奥行き方向の距離計測は , 両眼を含む平面と映像表示平面とが平行であると仮定して , 両眼間隔の距離から奥行きを推定しており , 頭部の前後移動と左右回転が区別できないため , 頭部姿勢の影響で映像が動いてしまい , 奥行き感が不正確になってしまう . 河原塚らはマーカレスで使用可能な ViewFrame[24] を提案している . 頭部の上下左右の位置は単眼カメラから色抽出により計測し , 頭部の奥行き距離は赤外線測距センサを用いている . このシステムは決まった位置を計測しているわけではないため , 視点位置がふらつくことが想定される . これらセンサ以外にも赤外線マーカを用いて頭部位置の計測を行っているものに , Djajadiningrat らの Cubby[25] がある . これは底面に 1 つ側面に 2 つの面をもつ 3 面スクリーンによる投影型 3 次元ディスプレイであり , Media Cube と同様に 3D CG を立体的に見ることが可能である .

このように , 従来のシステムでは頭部計測のためのデバイスやマーカの装着を強いられたり , 非接触・非装着に計測可能なものでも計測位置が不正確なものが多い . 一方 , 松本らはユーザの 3 次元顔モデルに基づいて , 精度の良い顔の 3 次元位置計測を実現している [26] . 本研究では , この計測システムをベースとし視点位置計測を行う . 詳しくは第 3.2 節にて述べる .

2.2.2 3 次元映像の提示方法の分類

3 次元映像を提示する方法としては , 3 次元モデルを用意したり , 3 次元復元を行い , モデルから提示する方法と 2 次元の画像を 3 次元モデルを生成することなく変換を行う方法の二つが挙げられる . 以下は , これらの代表的な手法について述べる .

モデルベースによるアプローチ 館らの Media Cube や Mulder らの AR システム , Pastoor らの視線入力可能な 3 次元ディスプレイ , Djajadiningrat らの Cubby などとは事前に用意された CG モデルを利用し , 運動視差の提示を行っている . Media

Cube は手に持つことが可能で，Cubby は据置型という違いはあるが，この二つは基本的に箱の中に置かれた物を鑑賞することができ，小型，大型のショーウィンドウのような使い方が想定される．また Mulder らの AR システムは机上での実物と仮想物体の操作などを想定している．Pastoor らの視線入力可能な 3 次元ディスプレイはコンピュータのワークスペースを立体的に利用したインタフェースを想定しており，グラフの 3 次元提示や，ドキュメントや Web ブラウザの空間的な配置，ワークスペース内でのオブジェクトウィンドウ（グラフやブラウザなど）の回転や移動などの操作を実装している．Kanade らは，多視点実画像からの 3 次元復元を行い，自由な視点での 3 次元映像生成を行っている [27][28]．ただし，3 次元復元を行う場合は計算コストが高いため完全なリアルタイムでの提示は困難であり，録画ビデオとしての利用が想定される．

イメージベースによるアプローチ 3 次元モデルを利用しない方法としては，以下の 3 つが挙げられる．

1. 各視点位置に対する画像配置やアクティブカメラ
2. モーフィング
3. 光線空間法

事前に撮影された複数視点の画像を各視点位置に配置したり，アクティブカメラを用いて任意の視点までカメラを動かし必要な画像を取得する方法としては，Lippman の研究 [29] や，Yanagida らの研究 [30]，財津らの研究 [31] などが挙げられる．

またビューモーフィングを用いた研究の代表例としては，Seitz らの研究があり，2 つの視点の画像からその中間視点の画像を生成している [32]．また石川らは複数台の全方位カメラ画像を用いてウォークスルー環境として利用可能な広範囲での立体映像生成を行っている [33]．

撮影された画像から 3 次元空間中を伝搬する光線情報として 5 自由度のパラメータ空間に記述し，光線情報から任意視点を復元する光線空間法としては，苗村らの研究や [34]，福嶋らの自由視点テレビ [35] などが挙げられる．また，石川

らは大規模空間での多視点映像を用いた自由視点映像提示を行っている [36] . 光線空間法を利用するためには密に撮影された複数視点の画像を必要とする . そのため , カメラ台数の制限などがある場合には光線空間パラメータが不足するため , 別視点での映像を復元することが出来ない .

本研究では , マニピュレータ先端に取り付けたカメラを用いてユーザ視点に対応する画像を取得し , モニタ面への射影を行うことで立体提示を行う .

2.3 運動視差 3 次元ディスプレイの評価

本システムは一般的に利用される両眼視差を用いず , 運動視差のみを提示することにより奥行き感の提示を行う . そのため両目で利用した場合には平面上の画像を両目で観察することになり両眼視差が立体感の不整合を与えてしまうことが想定される . そこでモデルベースシステムを用いて片目 , 両目における距離知覚実験を行い , 立体感の比較を行う . また 3 次元提示を行うために複数視点画像を撮影する場合には , 複数台カメラを利用したアプローチとアクティブカメラを利用したアプローチが想定される . 複数台カメラでは , どの程度のカメラ台数をどういった間隔で配置するかといった問題が存在し , またアクティブカメラでは , 複数台カメラと比較してカメラ移動の時間遅れが問題となる . そこでイメージベースシステムにおいては , 画像の空間的な提示間隔や時間遅れに対しての立体感の許容について評価実験を通して議論する .

2.4 本研究の位置付け

本研究では , モニタの前に座った一人のユーザを対象とし , カメラを用いた顔情報計測により装着を必要しない視点計測により運動視差の提示を行う . 運動視差提示方法には各視点位置に対する画像配置により実装を行い , 非接触な頭部位置計測に基づく運動視差 3 次元ディスプレイの有効性について検証を行う . また 3 次元ディスプレイの分野でたびたび無視される奥行き知覚実験を通して , システムの評価を行う .

2.5 本章の結論

本章では、運動視差に研究の成果と位置づけを明確にし、現在の問題点を指摘した。そして各応用分野における本研究の位置づけを従来研究との比較から明確にした。

3. 非接触な顔情報計測に基づく運動視差3次元ディスプレイの提案

3.1 本章の趣旨

本章では、非接触な顔情報計測に基づく運動視差提示3次元ディスプレイを実装する上で必要となるアルゴリズムについて議論する。まずベースとなる視点計測システムについて説明する。視点計測はカメラを用いて非接触に行われる。その後、ユーザ視点に対応した映像提示について説明する。固定モニタに対して映像を提示する場合、ユーザ視点位置によりモニタ画面が歪んで知覚される。これについて運動視差提示を想定して撮影された画像から歪みの確認と視点位置に応じた映像の変化を確認する。最後にこの歪みを補正するための射影変換について説明する。

3.2 顔情報計測に基づく視点位置計測

視点位置計測システムのベースとして、既存の顔トラッキングシステム [26] を用いた。このシステムは1組のステレオカメラを用いて、ユーザの3次元顔モデルを基に、顔の3次元位置・姿勢、視線ベクトルを計測することが可能である。位置・姿勢の計測には目の縁や口の縁などのテンプレートを利用したテンプレートマッチングで実現され、視線計測にはハフ変換が利用されている。処理速度はPentium4 2.53GHzを用いた場合、約10msでの処理を実現している。キャプチャによる遅れを含めると遅れ2フレーム以内でビデオレートでの処理が可能であり、それに同期して滑らかな映像描画を行うことが出来る。計測精度は画像のキャプチャサイズが 320×240 pixelsの時に、顔の位置計測精度が約2mm、顔の姿勢計測精度が約2deg、視線ベクトルの計測精度が約5deg程度である。既存システム [37][38] と比較して、マーカを付けることなく、非接触、受動的かつリアルタイムに頭部の3次元位置を計測することが可能であり、輝度値の正規化相互相関を用いているため照明の変化にも強く、また3次元表示を行うといった他の処理を行う余力も残されているため有効であると考えられる。

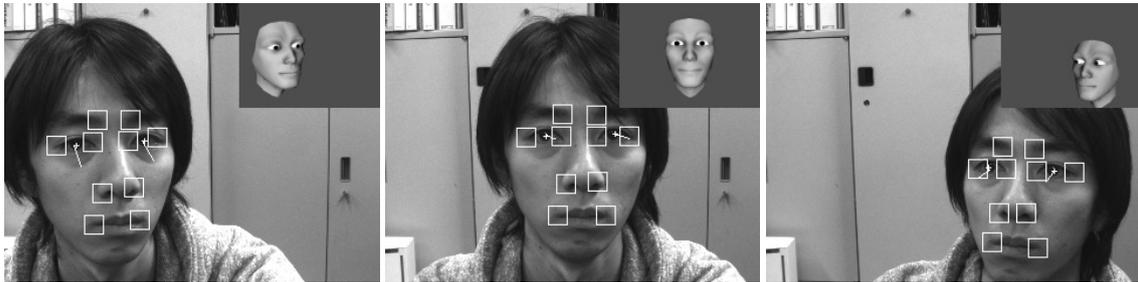


図 3 顔情報計測結果

本研究では，この計測結果として得られた両眼の目尻・目頭のテンプレートの 3 次元位置から左右の目の中間の 3 次元位置を求め，視点位置として利用した．図 3 に顔情報計測の結果を示す．

3.3 運動視差提示のための透視投影変換

3.3.1 視点位置により知覚される映像の違い

図 4 に実世界でモニタ枠越しに景色を見た時の見えの変化を示す．モニタ枠は液晶モニタから取り外した枠のみであり，左右には右への移動，上下には上への移動，斜めには前方への移動を表している．

右への移動を比較すると，正面の場合はユーザ視点から見たモニタ枠はほぼ長方形となっているのに対して，右から場合ユーザ視点から見たモニタ枠は歪んで知覚される．これはモニタが環境に固定して設置されているためで，ユーザの位置から遠い枠の左側では映像が縦に潰れて知覚されるため，モニタへの描画の際にはモニタの左側ほど大きく描画しておく必要がある．次に上への移動を比較すると上方から見ている方がより手前の景色が知覚されている．またユーザ視点位置から遠いモニタ枠の下側は左右に潰れて知覚されるため，左右移動と同様にモニタへの描画の際には補正して提示する必要がある．

前方への移動については，モニタに描画される画像の大きさは直感的な感覚とは逆の現象が起きている．通常，人は頭を前方に近付けた場合には物体がより大きく知覚されるためモニタにも大きく描画するように感じられるが実際には，前

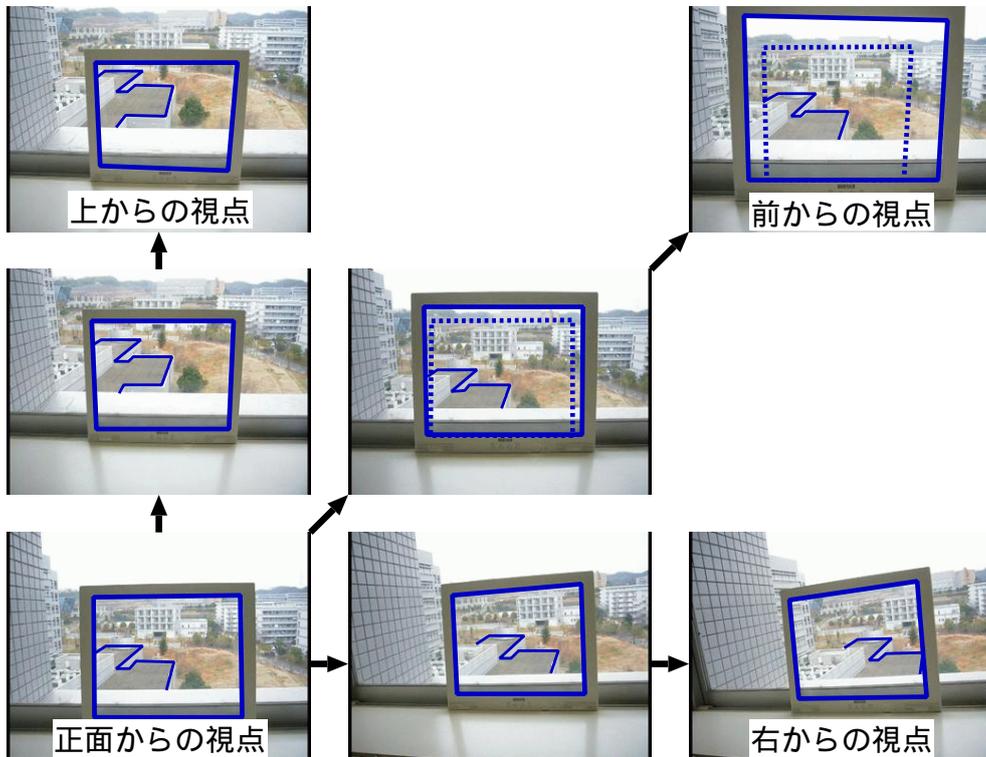


図 4 実世界でモニタ枠越しに景色を見た時の見えの変化

方に行くほどモニタを通して見える視野が広がるためモニタ内には多くの景色が現れる．結果として前方に行くほどモニタには小さく描画されることになる．ただしユーザにとっては頭部を前方に近付けた効果により画像が大きく知覚されることになり，この二つの効果の足し合わせにより物体の大きさが認識される．

3.3.2 固定モニタへの射影変換

射影変換 視点位置に対応した映像は，モデルベースシステムではCGモデルに対して，イメージベースシステムでは撮影された画像に対して図5のような非対称な四角錐形状の透視投影変換を行うことで生成される．モニターサイズを $w \times h$ とし，モニター面の中心を原点 O とした時の視点位置 V を (v_x, v_y, v_z) ， V から near クリップ面までの距離を n ，far クリップ面までの距離を f ，モニター面までの距離を d とする．また V から near クリップ面に下ろした垂線の足を P ，モニター面に下ろした垂線の足を Q と置く． P を原点とした near クリップ面左右の縁の x 座標を l, r とし， P を原点とした near クリップ面上下の縁の y 座標を b, t とする．同様に Q を原点としたモニタ面左右の縁の x 座標を l', r' とし， Q を原点としたモニタ面上下の縁の y 座標を b', t' とする．この時の透視投影行列は (1) となる．

$$\mathbf{R} = \begin{bmatrix} \frac{2n}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & -\frac{f+n}{f-n} & -\frac{2nf}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (1)$$

モニタサイズ w, h と視点位置 v_x, v_y, v_z により d と l, r, b, t を表すと (2) ~ (6) となる．

$$d = v_z \quad (2)$$

$$\begin{aligned} l &= l' \frac{n}{d} \\ &= \left(-\frac{w}{2} - v_x\right) \frac{n}{v_z} \end{aligned} \quad (3)$$

$$\begin{aligned} r &= r' \frac{n}{d} \\ &= \left(\frac{w}{2} - v_x\right) \frac{n}{v_z} \end{aligned} \quad (4)$$

$$\begin{aligned} b &= b' \frac{n}{d} \\ &= \left(-\frac{h}{2} - v_y\right) \frac{n}{v_z} \end{aligned} \quad (5)$$

$$\begin{aligned}
 t &= t' \frac{n}{d} \\
 &= \left(\frac{h}{2} - v_y \right) \frac{n}{v_z}
 \end{aligned} \tag{6}$$

よって最終的な射影行列はモニターサイズ w, h と視点位置 v_x, v_y, v_z を用い表すと (7) となる .

$$\mathbf{R} = \begin{bmatrix} \frac{v_z}{w} & 0 & -\frac{v_x}{w} & 0 \\ 0 & \frac{v_z}{h} & -\frac{v_y}{h} & 0 \\ 0 & 0 & -\frac{f+n}{f-n} & -\frac{2nf}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{7}$$

near クリップ面の位置をモニター面にした場合 ($n = d$) , 物体はモニターより奥に存在するよう描画され , far クリップ面の位置をモニター面にした場合 ($f = d$) , モニターの手前に存在するよう描画される . 運動視差はこの映像を逐次 , ユーザに提示することで得られる .

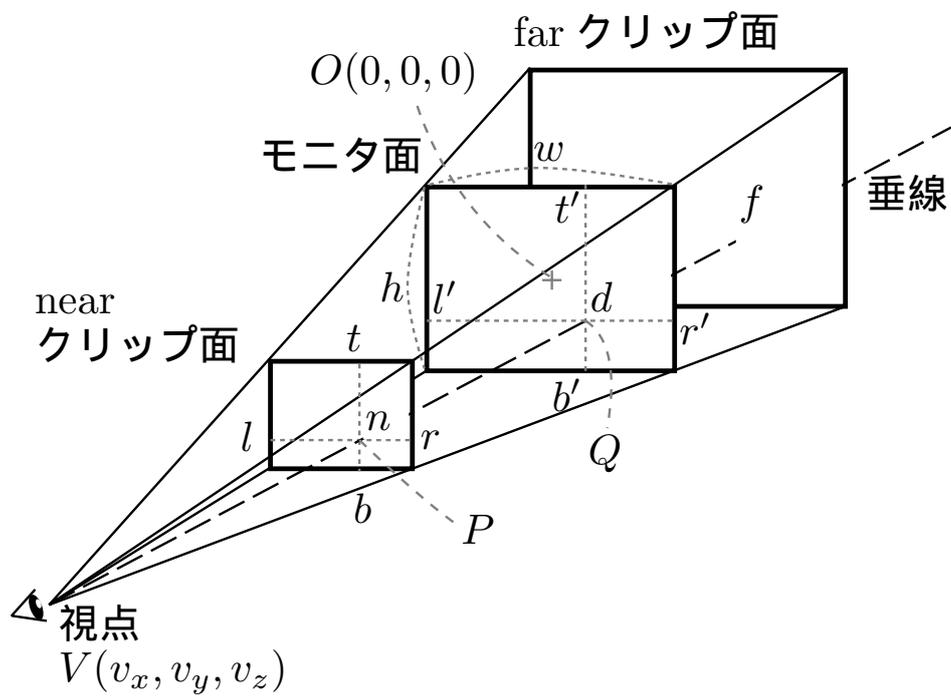


図 5 非対称な透視変換

イメージベースにおける画像取得と射影変換 実画像を用いる場合には，撮影されたキャプチャ画像を視点位置に応じて射影変換し，ユーザへ提示する．図 6 に画像取得と射影の概要を示す．カメラにより計測された視点位置と同じ位置からのキャプチャ画像を取得する．この時，キャプチャ画像はカメラの位置から焦点距離だけ離れた位置に配置される．この時の画像の姿勢は，撮影を行ったカメラの姿勢に対応する．カメラ光軸は撮影画像平面と直交し，画像中心を通る．この画像に対して式 (7) の射影変換を行い，射影画像をモニタに表示させることで，ユーザ視点に対応した画像を提示する．この一連の処理を逐次行うことで運動視差の提示が行われる．

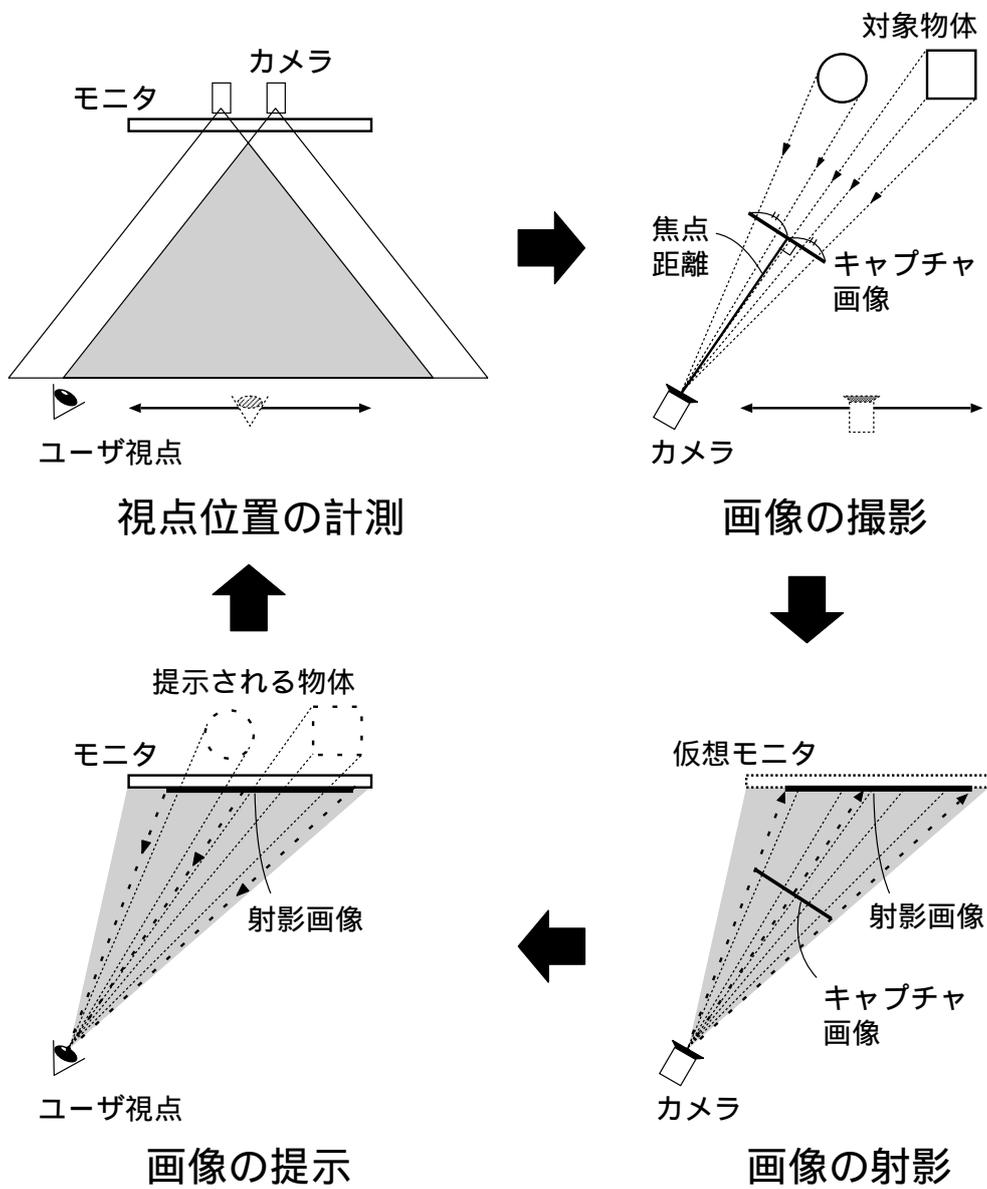


図 6 視点に対応した画像取得と射影変換

3.4 本章の結論

本章では、非接触な顔情報計測に基づく運動視差提示3次元ディスプレイの提案を行い、実装する上で必要となる視点位置計測方法と射影変換について説明をした。固定モニタに対して映像を提示する場合に、ユーザの視点位置に応じて知覚されるモニタ枠の形状が変わることを確認し、モニタへの映像提示を行う上で視点位置に対応した補正を行う必要があることを示した。最後に映像補正のための射影変換について説明をした。

4. 非接触な顔情報計測に基づく運動視差3次元ディスプレイの実装

4.1 本章の趣旨

本章では、モデルベースやイメージベースにおける運動視差3次元ディスプレイの実装について述べる。ユーザの視点位置はステレオカメラにより計測され視点位置に対応した画像を生成しユーザに提示する。モデルベースの場合は読み込んだ3D CGモデルを、イメージベースの場合にはある視点の画像をそれぞれ射影変換することでユーザ位置において整合性のとれた画像を生成する。イメージベースによるシステムでは、事前に収集した多視点画像を利用するオフラインシステムとユーザの視点移動に追従してカメラを動作させるオンラインシステムについて説明する。多視点画像の撮影は6軸マニピュレータの先端に取り付けられたカメラを用いて行う。

4.2 実装1：モデルベースによる運動視差提示3次元ディスプレイ

本システムは視点位置計測システムと映像提示システムとで構成されている。視点位置はステレオカメラを用いて非接触、非拘束に計測され、視点位置に応じたCG画像は透視投影変換を行うことで生成される。このCG画像は逐次、平面ディスプレイに提示され、ユーザは頭を動かしながら見ることで画像を立体的に知覚することが可能である。また頭部計測範囲を拡大させるため、オプションとして回転ステージが利用可能であるが、回転ステージを用いなくとも運動視差の提示は可能である。以下に詳細を述べる。

4.2.1 モデルベース3次元ディスプレイのシステム構成

ハードウェア構成を図7に示す。ディスプレイデバイスとして通常の平面液晶ディスプレイを用いている。ユーザの視点位置は1組のステレオカメラ(Point Grey社製IEEE1394カメラFlea,2台)から計測される。またオプションとして利

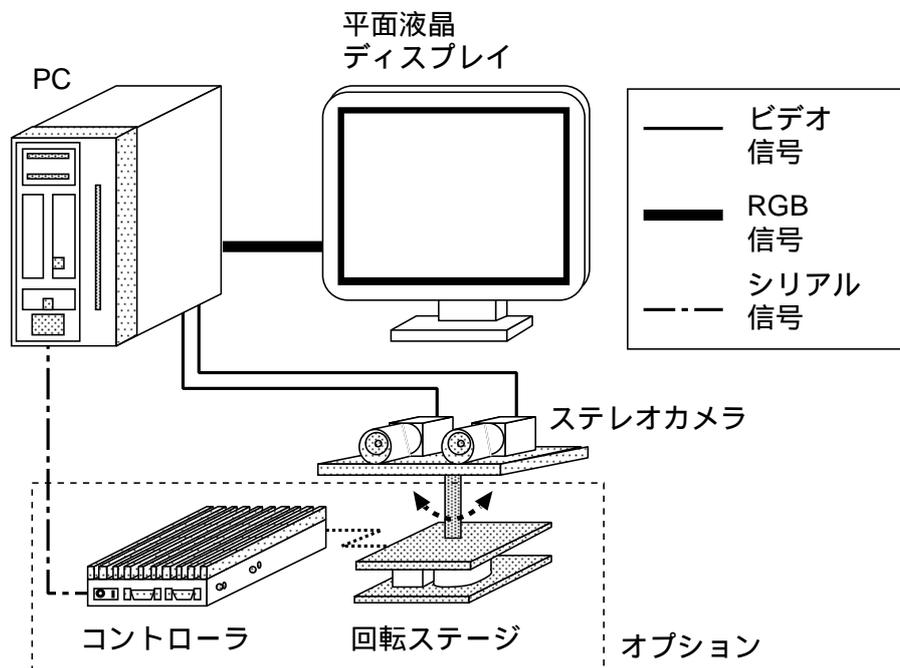


図 7 モデルベース 3次元ディスプレイのハードウェア構成図

用可能な回転ステージにはステッピングモータ制御のシグマ光機製回転ステージ SKIDS-60YAW を用いており，制御コントローラ CSG-602R から RS232C を用いて制御される．実際に運動視差を提示した様子を図 8，図 9，図 10 に示す．視点位置計測が約 10ms で終了しており，ポリゴン数の少ない 3D モデルであれば，キャプチャ時間の 30ms 以内で映像の提示が可能である．

4.2.2 回転ステージによる計測範囲拡大

回転ステージは頭部計測範囲拡大のためのオプションとして利用する．頭部位置視線計測により得られた 3次元位置をキャプチャ画像に投影した点が画像中心に向かう方向に回転させる．現在の動作は速度 $\pm 15\text{deg/s}$ による bang-bang 制御で行っており，頭部が画像中心に存在する付近（画像幅の $\pm 1/4$ ）は停止させている．回転ステージ SKIDS-60YAW の最大回転速度は約 60deg/s であるが，回転が速過ぎる場合には画像のぶれが大きくなるため，本システムでは比較的低速な

頭部の動き（モニタ面から 60～70cm 離れた位置で左右に 15cm/s で動ける範囲）を追跡することを想定している。

最終的な視点位置は頭部位置計測結果に対して回転ステージから得られた回転量の回転補正とスクリーン・カメラ間の位置姿勢分の並進回転補正により得られる。回転ステージの回転量が 0deg の時にカメラが正面を向いているとし、この時のモニタ原点に対するカメラの位置姿勢を t_{cam} 、 R_{cam} とし、回転ステージの姿勢を R_{stage} とすると、変換行列 $D_{cam2disp}$ は式 (8) のように表される。ただし、1 軸の回転ステージを用いているため、 R_{stage} は Y 軸回りの回転のみである。

$$D_{cam2disp} = \begin{bmatrix} R_{Stage}R_{cam} & t_{cam} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \quad (8)$$

$$\mathbf{0}_3 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$$

回転ステージを利用した場合は回転ステージの制御および、それに同期した映像描画も 10Hz となるため、ビデオレートでの描画と比較すると滑らかではない描画となっている。しかし、左右の利用可能範囲はステレオカメラの共通視野である約 $\pm 10\text{deg}$ から平面ディスプレイの視野角を超える $\pm 90\text{deg}$ まで広げることが可能であることが確認された。回転ステージの使用については用途に応じた利用が適切であり、比較的静的で広い環境（例えば窓から見える山などの景色）を眺める場合などには有効であると考えられる。

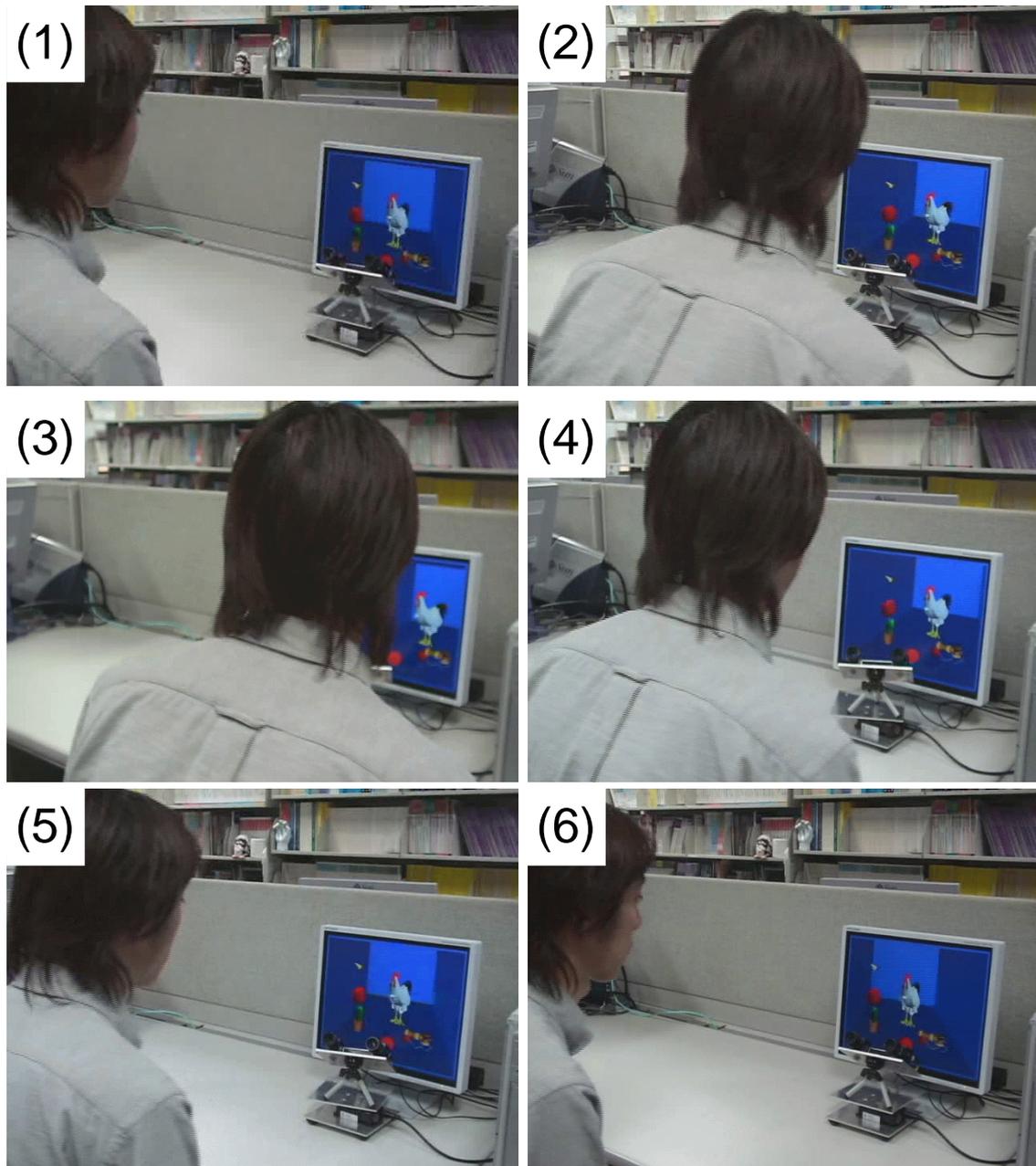


図 8 モデルベース 3次元ディスプレイの運動視差提示の様子

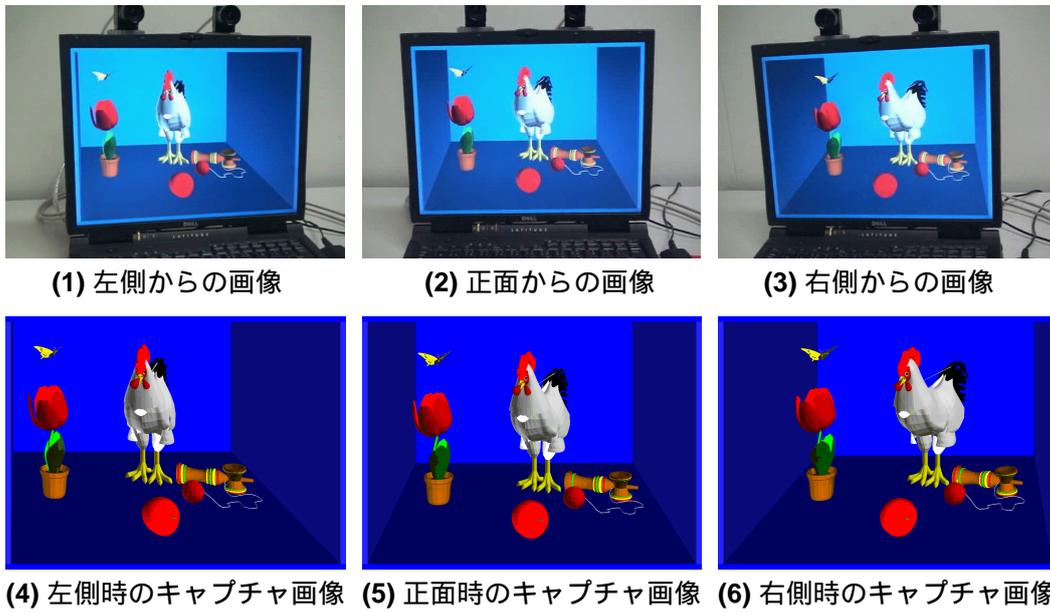


図 9 モデルベース 3 次元ディスプレイの回転ステージを用いない運動視差提示結果

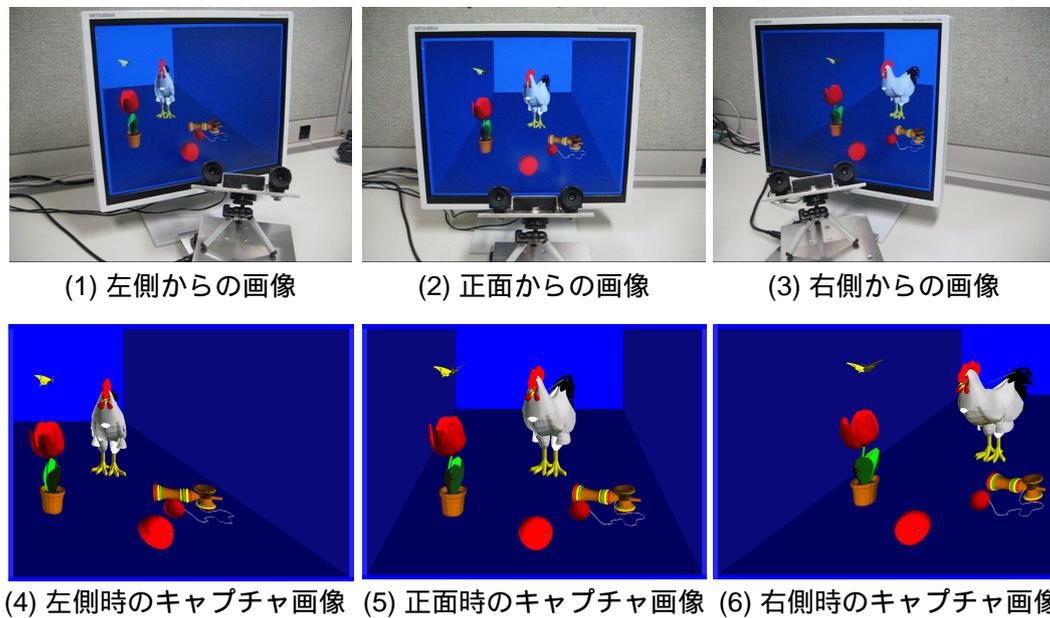


図 10 モデルベース 3 次元ディスプレイの回転ステージを用いた運動視差提示結果

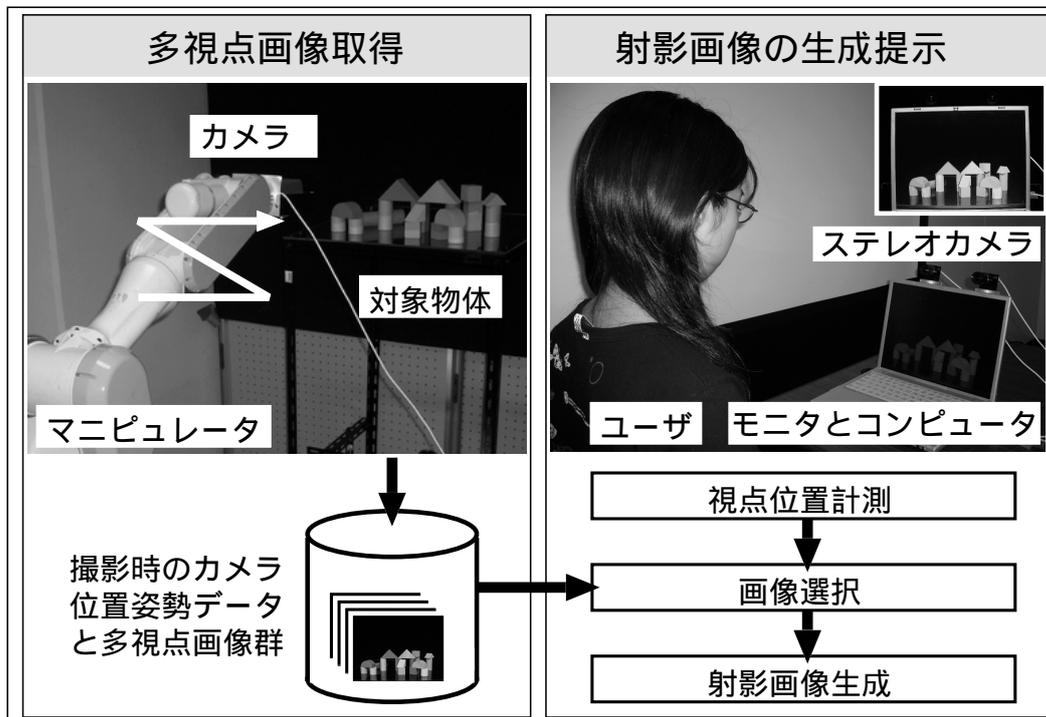


図 11 画像データベースを用いた運動視差提示 3 次元ディスプレイの概要

4.3 実装 2：多視点画像データベースを用いたイメージベース運動視差提示 3 次元ディスプレイ

図 11 にマニピュレータによる画像撮影の様子と運動視差に基づく実画像 3 次元ディスプレイの外観を示す。複数枚画像はマニピュレータに取り付けられたカメラにより事前に取得され、蓄積された画像群を視点位置に対応した画像を選択し、射影変換を行うことで提示される。

4.3.1 マニピュレータによる多視点実画像の撮像

提示画像の撮影は 6 自由度マニピュレータに取り付けられたカメラを用いて行う。複数枚の画像群は上下、左右、前後などいくつかの軸方向に対して直線的に制御しながら撮影される。また提示するモニタは通常の平面モニタを対象として

おり、効率的な撮影のため撮影対象をカメラの注目点として、常にカメラの光軸が注目点を貫くように画像の撮影を行う。取得された画像は参照点を基準としたカメラの位置、姿勢と共に保存される。

カメラは IEEE1394 カメラ (SONY DFW-VL500) を使用し、カメラキャリブレーションにより事前に内部パラメータを求め、レンズの歪み補正を行った。

4.3.2 多視点実画像からの運動視差提示

複数枚の多視点画像は HDD からの画像探索時間短縮のために、図 12 のようにカメラ位置に応じて上下、左右、前後など格子状に整列させる。これらの画像の中から位置がユーザ視点位置に一番近い画像が選択され、逐次 HDD から画像を読み込み射影、描画される。

左右に 1mm 刻み $\pm 150\text{mm}$ の 301 枚、上下に 10mm 刻み 10 枚の合計 3010 枚の画像を用いて運動視差提示を行っている様子を図 13、図 14 に示す。この時の処理時間を表 1 に示す。処理時間は CPU Core Duo 1.66GHz、メモリ 2GB のコンピュータにより測定した。前処理として画像整列のための起動時間が必要となるが、画像そのものの読込は行わないため、3010 枚の画像に対しても約 90ms で処理が完了する。起動後は逐次、画像を HDD から読み込む必要があるが、視点位置計測の処理時間と合わせてもキャプチャ時間の 30ms 以内に完了している。

表 1 多視点画像データベースを用いたシステムの処理時間。画像サイズ 640 × 480

項目	処理時間 [ms]
前処理 (画像の整列)	$0.03 \times \text{画像枚数}$
視点位置計測	10
HDD からの画像読込	10
射影と描画	5
合計処理時間	25

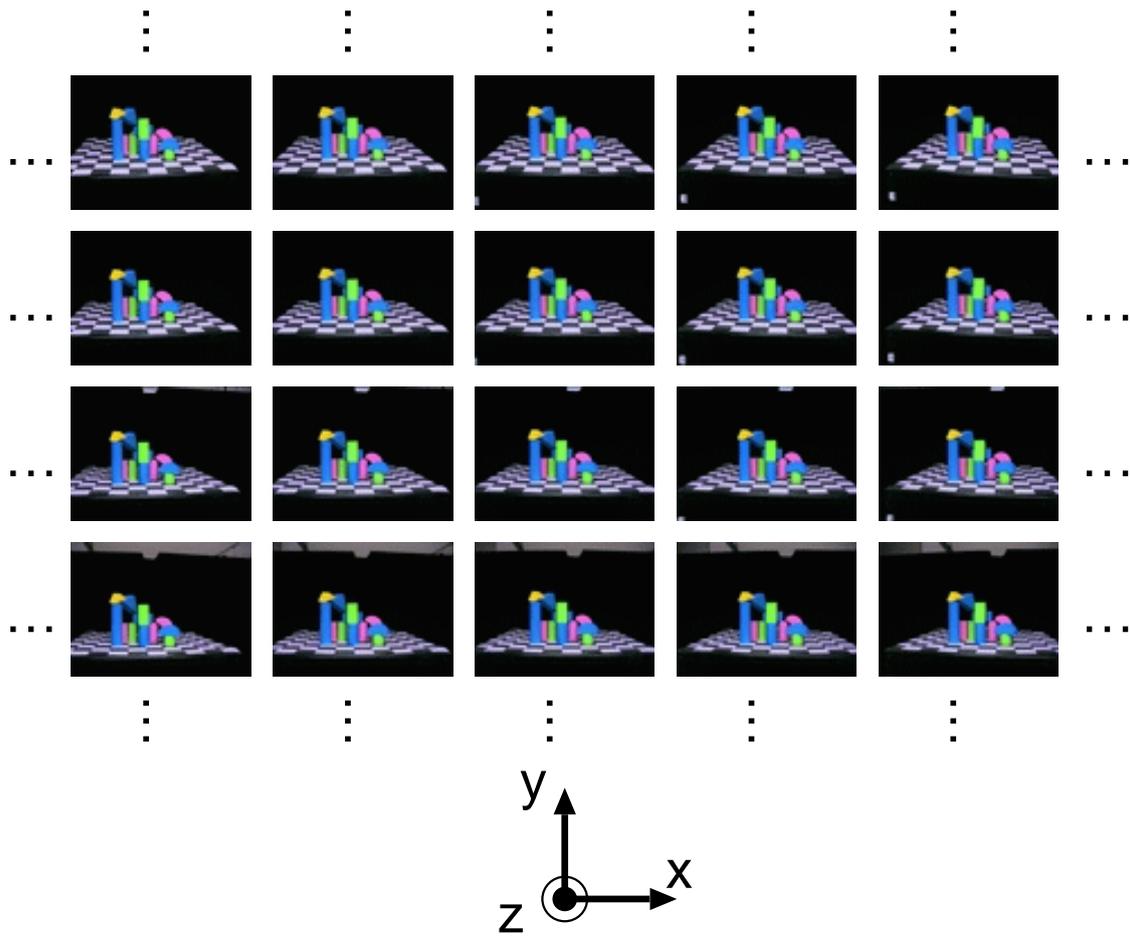


図 12 多視点画像データベースの構造

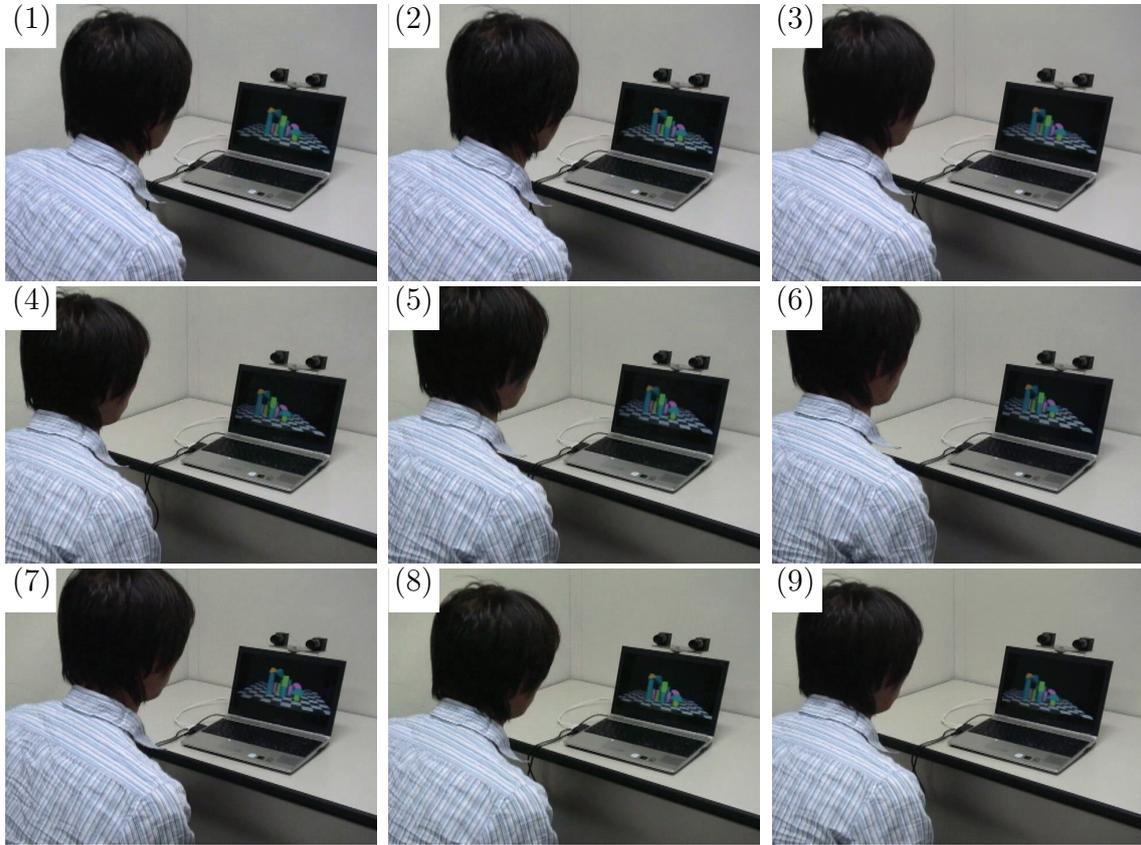


図 13 多視点画像データベースを用いた運動視差提示の様子



図 14 多視点画像データベースを用いた運動視差提示の結果

4.4 実装3: オンラインでのイメージベース運動視差提示3次元ディスプレイ

図 15 にオンラインによる運動視差提示3次元ディスプレイの処理の流れを示す。本システムではユーザの視点位置を計測し、計測された視点に追従するようにカメラの制御を行う。カメラの姿勢は、撮影対象を常に向くように位置に応じて決定される。その後現在のカメラの位置姿勢を逐次ユーザ側に伝送され、カメラからの画像の取得とレンズ歪みの補正を行う。最後にカメラ位置姿勢に対する射影変換により生成された画像をユーザへ提示し、これら一連の流れを逐次行うことで運動視差の提示を行う。

映像提示の様子を図 16 に、映像提示を行った結果の4秒間隔のスナップショットを図 17 に示す。この時のユーザ頭部の移動とマニピュレータの移動の結果は図 18, 図 19, 図 20 のようになり平均動作遅れ時間は668msであった。オンラインシステムの処理時間を表 2 に示す。

オンラインシステムでは全体で約728msの映像提示の遅れが生じているが、データベース画像データベースでは不可能である動物体の実画像立体提示が可能となる。時間遅れの多くはマニピュレータの動作遅れであり、動作速度の向上や応答性を改善が必要であると考えられる。実際に立体感の提示を行う上で、映像の提示遅れの許容については第5.3.3節において評価、検討を行うが、オンラインのシステムでは、映像遅延の問題を取り除いて立体感の評価を行うことができない。しかし、多視点画像データベースのシステムでは、映像提示がキャプチャ時間の33ms以内で実現できており、ほぼ遅れなく提示可能である。そこで、イメージベースによる立体提示の評価については、多視点画像データベースによるシステムを利用する。

表 2 処理時間

項目	処理時間 [ms]
視点位置計測	10
視点位置送信	5
マニピュレータ動作遅れ	668
カメラ位置姿勢送信	5
画像キャプチャ	5
レンズ歪み補正	30
射影と描画	5
合計	728

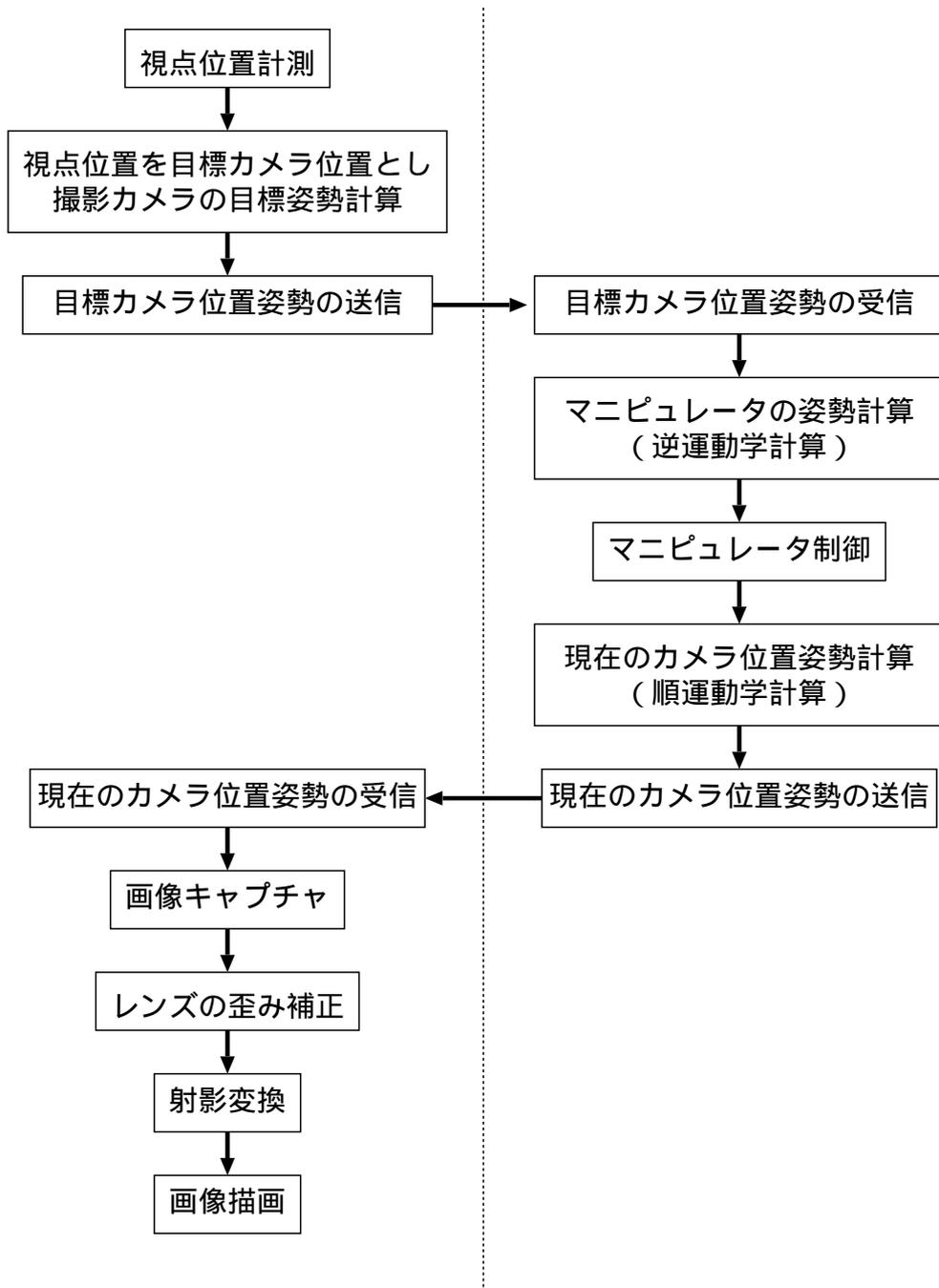


図 15 オンラインによる運動視差提示 3次元ディスプレイの処理の流れ



図 16 オンラインによる運動視差提示の様子

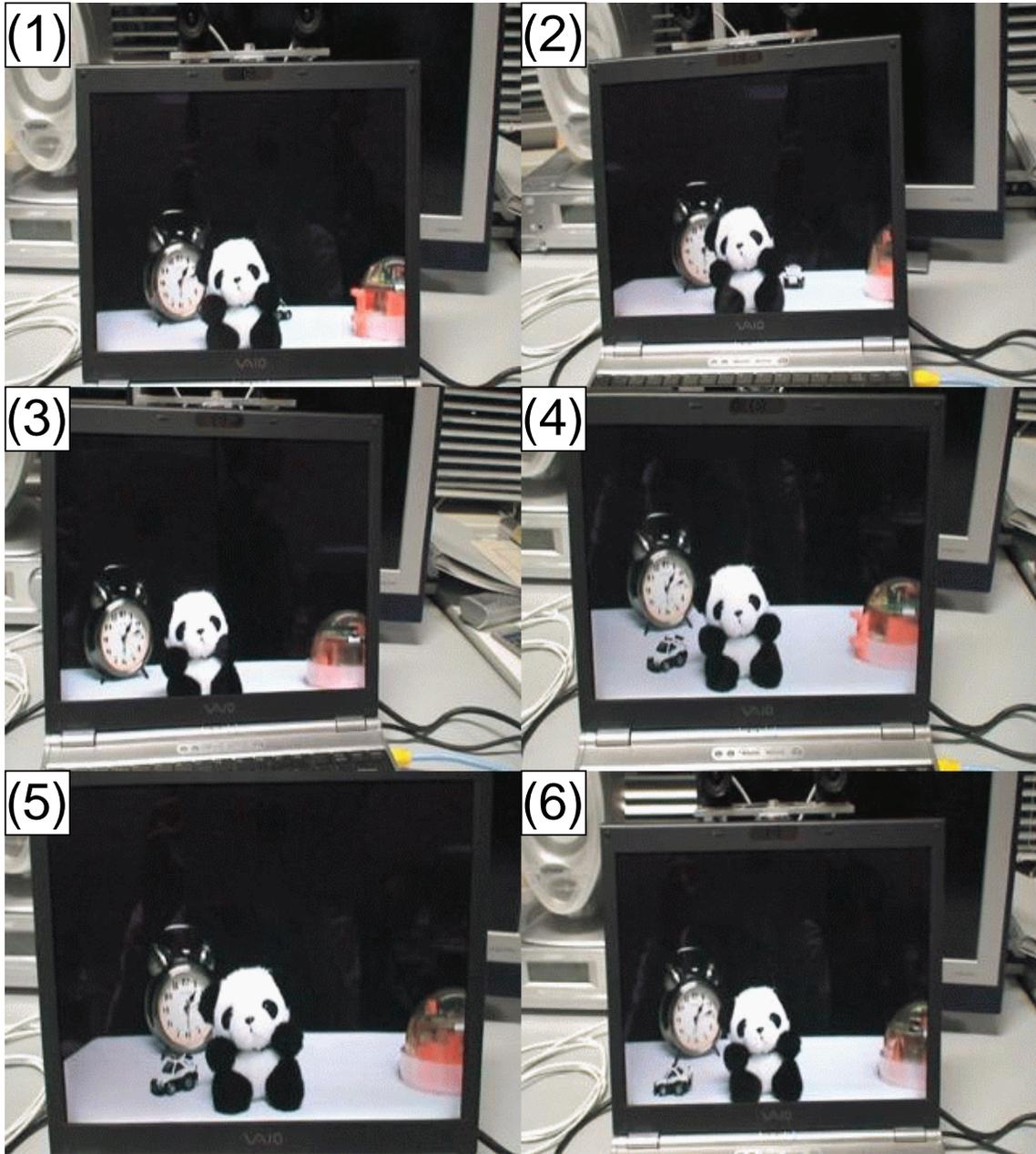


図 17 オンラインによる運動視差提示結果 (4 秒間隔)

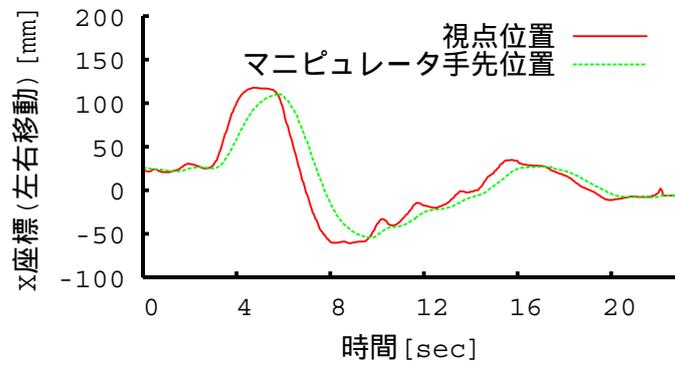


図 18 マニピュレータ左右移動の遅れ

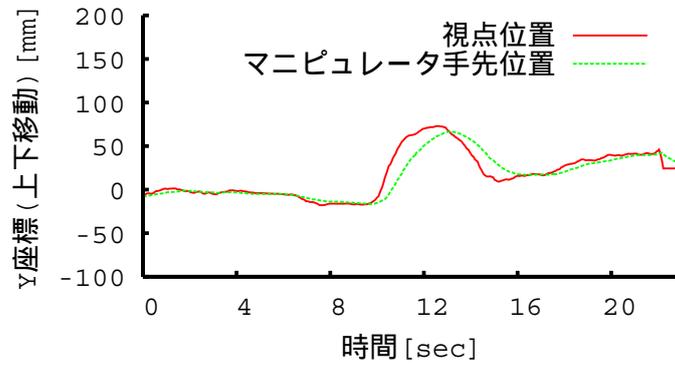


図 19 マニピュレータ上下移動の遅れ

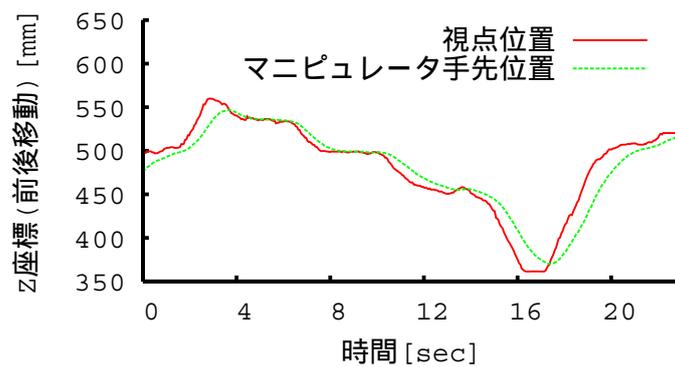


図 20 マニピュレータ前後移動の遅れ



図 21 トラッキング用マーカを取り付けたマニピュレータ

4.5 イメージベースによる運動視差 3 次元ディスプレイの提示画像 評価

射影による画像の変換の正当性を確認するために、マニピュレータを用いて撮影した画像とユーザ視点の位置から撮影される画像の類似性を確認した。多視点画像の蓄積のために、マニピュレータを水平に 1mm 間隔で 300mm 動かし画像の撮影を行い、合計 301 枚の画像を保存した。この時、モニタからカメラまでの距離は 450mm である。ユーザ視点からの画像を取得するためにマニピュレータ先端に取り付けたビデオカメラにマーカを貼り付け、カメラの位置計測を行いながら撮影を行った。図 21 にマーカを取り付けたマニピュレータを示す。頭部の動きはマニピュレータを用いて再現している。この時、焦点距離は 760mm で画像のキャプチャサイズは 640×480 であった。

実験結果を図 22 に示す。事前に撮影したキャプチャ画像 (a) とユーザ視点から撮影された画像 (c) を比べると物体の位置やサイズ、オクルージョンがほぼ正しく提示できていることが分かる。よってユーザへの正確な運動視差提示が行えていることを確認した。

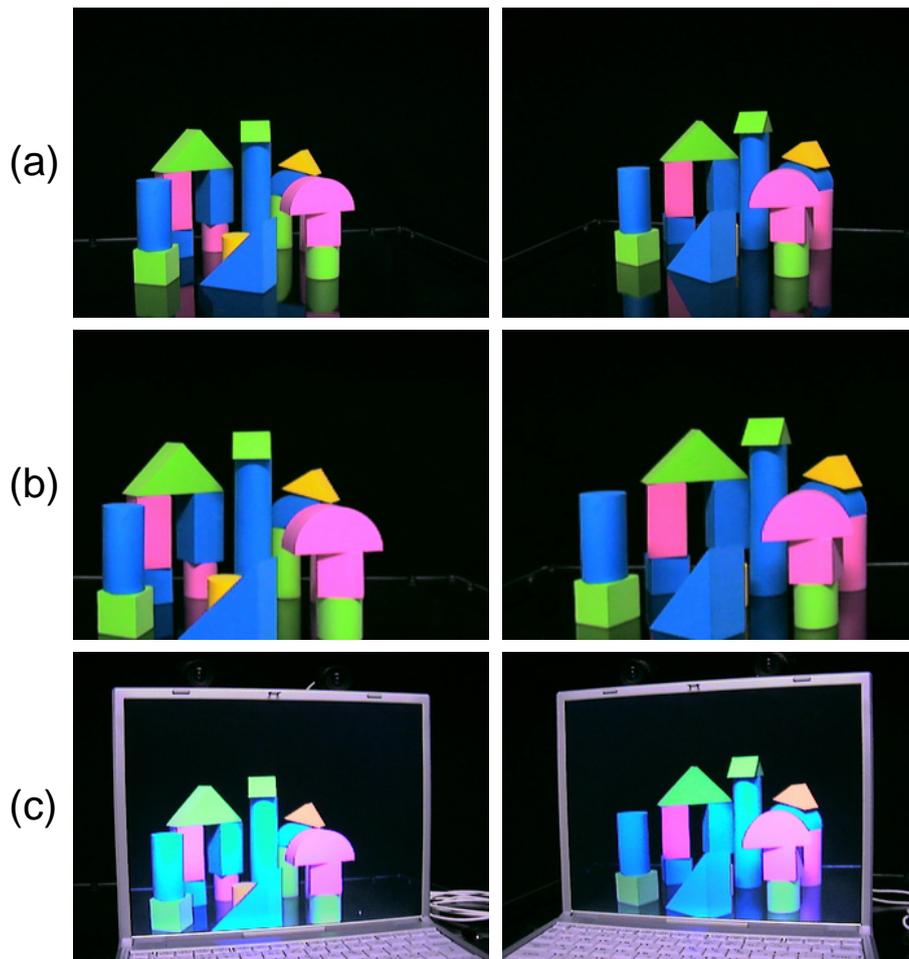


図 22 提示画像の整合性の確認実験結果 (a) キャプチャ画像, (b) 生成画像 (射影画像), (c) ユーザの視点位置から観測される画像

4.6 本章の結論

本論文では非接触に利用可能な運動視差提示3次元ディスプレイとして、モデルベースのシステム、多視点画像データベースを利用したイメージベースのシステム、マニピュレータをオンラインで動作させたイメージベースシステムを実装した。視点位置はステレオカメラを用いて計測され、ユーザは何も取り付けることなく、頭を動かしながら通常の平面ディスプレイの映像を見ることで画像を立体的に知覚することが可能である。モデルベースのシステムでは、3次元モデルデータを用意できる場合などでの利用が可能である。またインターネットを介した新製品紹介の3Dビューなどでは多視点画像の切替えから立体提示が行われており、多視点画像データベースのシステムを応用することにより実現が可能である。多視点画像データベースのシステムでは動きのない物体の映像提示をほぼ映像遅れなく行うことができる。オンラインシステムでは、約728msという映像遅れの問題は存在するが動物体に対する立体提示が可能である。遅れ発生のほぼ9割はマニピュレータの動作遅れによるものであり、カメラ制御の応答性向上により大幅な遅れの改善が見込まれる。

最後に、多視点画像データベースを利用したシステムを用いて提示画像の整合性について検証を行い、正確な運動視差の提示が行えていることを確認した。

5. 奥行き知覚に基づく評価実験

5.1 本章の趣旨

本章では、奥行き知覚実験を通して実装したシステムの評価を行う。

本システムは一般的に利用される両眼視差を用いず、運動視差のみを提示することにより奥行き感の提示を行っている。そのため両目で利用した場合には平面上の画像を両目で観察することになり両眼視差が立体感の不整合を与えてしまうことが想定される。そこで、モデルベースシステムを用いて片目、両目における絶対距離知覚実験を行い、奥行き感の比較実験を行った。

またイメージベースの多視点画像データベースシステムを用いた遅れのない画像提示により、画像の空間的な提示間隔を変化させた場合と擬似的に映像遅延を生じさせた場合における奥行き感の比較を物体の配置精度の客観評価とユーザアンケートによる主観評価により確認した。

5.2 モデルベースによる運動視差 3次元ディスプレイの評価

5.2.1 運動視差のみによる絶対距離知覚

実験方法 モデルベースによる運動視差 3次元ディスプレイシステムを用いて運動立体視による絶対距離知覚実験を行った。本実験と、5.2.2節の実験については回転ステージを使用していない。被験者は20代男性の大学院生7名で、頭を動かしながら様々な距離に置かれた直方体を観察してもらった。被験者には事前に紐を渡してあり、被験者が感じたモニターから直方体前面までの距離を紐の長さで表現してもらい、実験者は紐の長さをメジャーで測定し、距離の計測を行った。この方法は、心理学における大塚らの運動視差による奥行き知覚の実験 [39] において被験者が知覚した距離を別の物体の長さによって表現していることを参考にした。

実験で描画した画像は、黄色の直方体を真っ暗な空間に浮かべた状態であり、図23のようになっている。CG内の照明については知覚対象が明るく見える任意の位置に配置しており、陰影は描画していない。提示した物体の大きさ、距離を

図 24 に示す．高さはモニタ面の下端と同じ高さであり，左右方向については被験者がモニタ正面に座った時にモニタ中央に見えるよう配置されている．距離が近い順に Near 物体，Middle 物体，Far 物体とした．提示の順番は線形合同法による乱数で並べ替え，3 種類の直方体を各 1 回ずつランダムに提示した．これを 1 試行として，全 6 試行の実験を行った．全 6 試行のうち第 1 試行は練習として分析から除外している．また各描画を切替える間には真っ暗な画面を表示させた．実験は片眼・両眼それぞれについて行い，実験中，回りの環境を見ることや観察時間には制限を与えなかった．ここでは実際に提示した距離を観察距離，実際に知覚された距離を知覚距離と定義する．

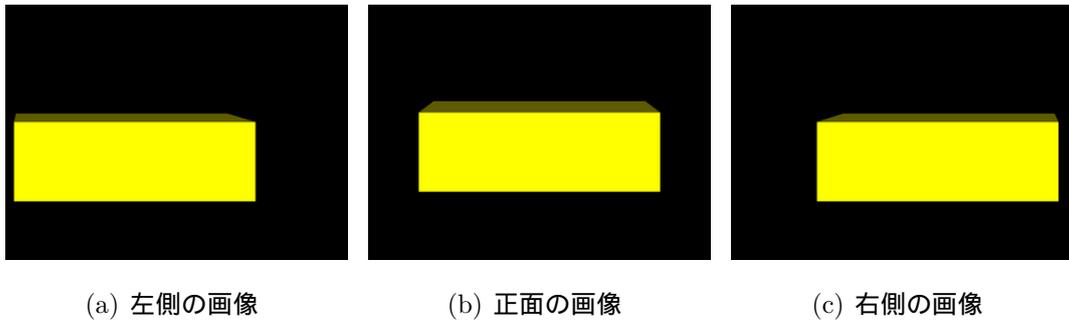


図 23 絶対距離知覚実験での描画画像

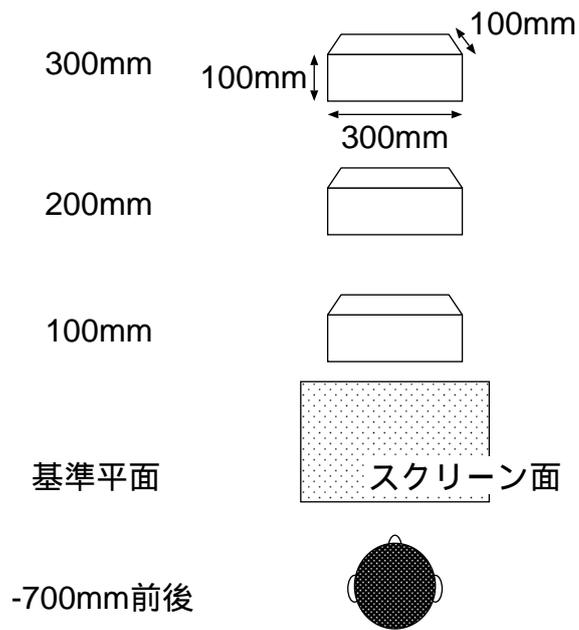


図 24 実験で使用した視覚対象

実験結果と考察 両眼の場合において被験者の1人が立体知覚をすることができず、実験を中断した。これは運動視差などによる奥行き手がかりの影響を越えて両眼視差や輻輳、遠近調節の違和感が大きく影響したものと考えられる。このような運動視差のみを用いて両眼視差などの影響を受けた場合に定位置に知覚されないというケースは西田らによっても報告されている [40]。

両眼で立体知覚が行えなかった被験者を除いた6人全体の片眼における知覚距離の結果を図25(a)に、両眼における知覚距離の結果を図25(b)に示す。棒グラフが平均値、誤差棒が標準偏差である。知覚距離100, 200, 300mmに引かれた横線は観察距離である。表3には、観察距離により正規化した知覚距離の誤差を示す。正規化誤差とは誤差を観察距離で割った値である。片眼両眼どちらの場合も観察距離が大きくなるにつれて知覚距離も相対的に大きくなった。各知覚距離間の平均の分散分析からも有意差が確認され、距離の違いを知覚できていることが確認できた。しかし、Near, Middle, Far 全てにおいて知覚距離が観察距離よりも小さくなる傾向が見られた。さらに、観察距離が Near, Middle, Far と距離が遠くなるにつれて誤差が大きくなった。これは、表3からも確認することができ、遠くなるにつれて奥行き知覚が不正確になっていることを示している。これらは焦点の遠近調節との違和感からくるものと考えられる。対象はスクリーンから離れた位置に配置されているが焦点はスクリーン上に合ってしまうため感覚の差異が生じてしまう。さらに距離が遠くなった場合には、焦点の合う距離と観察距離との差がより大きくなるために誤差も大きくなるものと考えられる。この結果は Hayashibe らによる Virtual Reality(以下, VR) 空間における絶対的奥行き距離知覚は、現実の奥行き知覚よりも過小評価されるという報告 [41] とも一致する。

これらの事から、本システムを利用する場合には、例えばインターネットでの Web ショッピングなどの絶対距離知覚が必要ない対象に対して利用することが適していることが示唆された。

この実験においては片眼、両眼の違いによって知覚距離に差がでると予想していたが、被験者の一人が両眼の場合に立体視ができなかった以外には、片眼の場合において有意差は確認されなかった。また、Far 物体においては両眼の方が正確に知覚されるなど片眼両眼の条件においては必ずしも片眼の場合が有意である

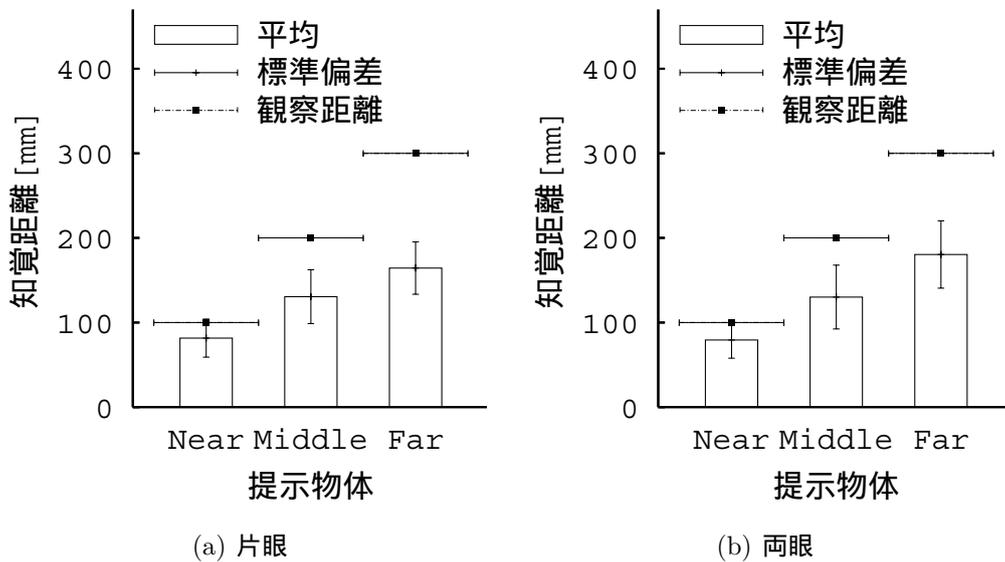


図 25 片眼両眼における知覚距離の平均

表 3 片眼両眼における知覚距離の正規化誤差

	片眼	両眼
Near	-0.18	-0.21
Middle	-0.35	-0.35
Far	-0.45	-0.40

という結果にはならなかった．むしろ片眼で利用した場合には両眼で利用した場合よりもより目の疲れが感じられたとの被験者からの意見も得られた．

5.2.2 絵画的要因との統合による絶対距離知覚

商品展示といったショーウィンドウ内にある商品の閲覧や居住空間設計，展示といったショールームの閲覧など実際の利用例を考えた場合，観察物体一つのみが提示されるだけでなく，観察対象の回りを壁が囲んでいたり，複数の物体が置かれた状態というのは十分想定される．そこで，単純な物体の回りを壁に囲ま

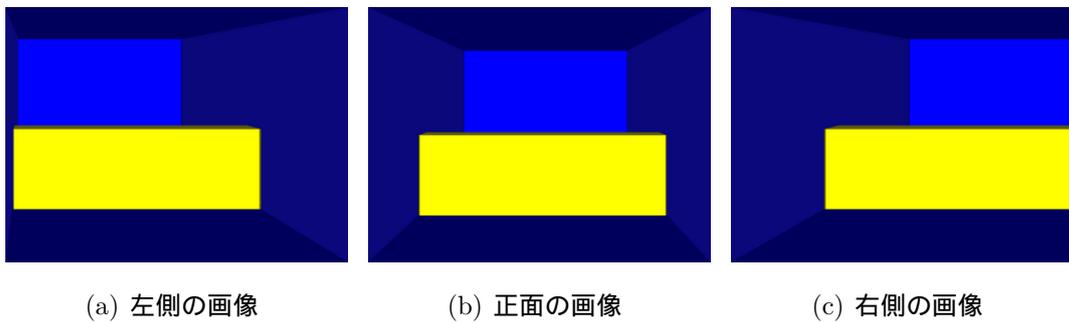


図 26 壁の描画と合わせた絶対距離知覚実験での描画画像

表 4 壁あり・なしでの知覚距離の正規化誤差

	片眼		両眼	
	壁なし	壁あり	壁なし	壁あり
Near	-0.18	-0.11	-0.21	-0.14
Middle	-0.35	-0.28	-0.35	-0.26
Far	-0.45	-0.39	-0.40	-0.36

れた空間における立体提示を行った。こういった他の物体が表示されている場合には視野内の高さの違いや重なり合いといった奥行き知覚手掛かりの絵画的要因の影響により立体感の向上が見込める。そこで3次元知覚の絵画的要因との統合の影響を実験を通して検討、議論する。

実験方法 実験方法は、5.2.1節と同様であり、描画対象としてはスクリーンの幅、高さと同じ長さの辺を持つ平面（以降「壁」とする）をスクリーンの側面から奥に伸びるように一緒に描画した場合について実験を行なった。提示物体の直方体、CG内の照明や陰影についても5.2.1節の実験と同様であり、物体は青い壁に囲まれた空間に置かれている状態である。（図26）。

実験結果と考察 図27(a)に片眼壁ありでの知覚距離の結果を、図27(b)に両眼壁ありでの知覚距離の結果を示す。また表4に観察距離で正規化した知覚距離の

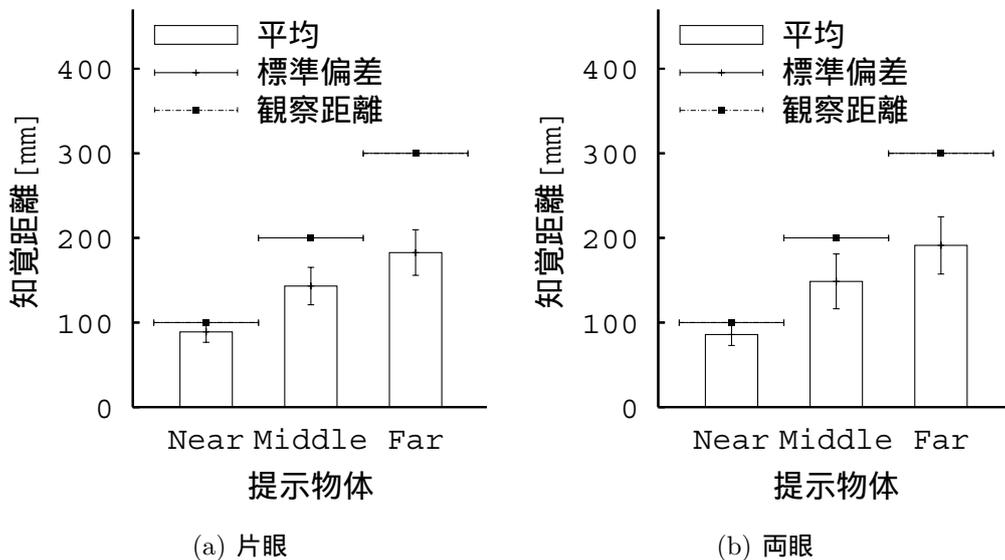


図 27 壁ありでの知覚距離の平均

誤差を示す．壁を共に描画した場合においても，知覚距離が観察距離よりも小さくなり，観察距離が大きくなるにつれて誤差が大きくなるという壁なしの場合と同様の傾向が得られた．しかし，表 4 より壁を共に描画した場合には全ての条件において知覚距離の誤差が小さくなった．Meng らによれば，自然空間内と同様にコンピュータでシミュレートした自然空間においても，知覚対象とそれが置かれた物体，そしてそれらが存在する地面との間の知覚的構造関係が重要であることが示されている [42]．さらに，3次元知覚の絵画的要因の統合においては加算的効果があり，各奥行き手がかりは単独モジュールで処理され [43]，それらが手がかりごとに重みづけられて加算され，最終的な奥行き・立体効果を生み出す [44] ことが報告されており，知覚手がかりの統合による加算の影響を受けたものと考えられる．

片眼両眼の影響と壁の有無の影響についての個人差を検討するため各被験者毎の知覚距離を 2 要因分散分析 (Two-way ANOVA) により比較した．各被験者の知覚距離の誤差を表 5 に，分散分析を行った結果を表 6 に示す．表 6 内の「両眼」や「片眼」，「有り」，「無し」は誤差が小さい項目であり，(**) は有意水準 1%，

(*) は有意水準 5% で有意差が認められた。(+) は有意水準 10% で有意傾向があり、(n.s.) は有意差が認められなかった項目である。相互作用については相互作用が認められた項目に「」が付けられている。被験者 G は両眼壁なしの条件で立体知覚が行えなかった被験者であり、表 5 では立体知覚が行えた壁ありの条件のみ誤差を示している。また検定については被験者 G のデータのみ壁有り条件での片眼両眼の t 検定と片眼条件での壁有無の t 検定により行っている。

被験者 G のように壁無しでは両眼において立体知覚できなかった被験者に対しても、壁を共に描画することで立体知覚が行われた。また、全被験者において絵画的要因と組み合わせることで両眼での利用時の誤差が小さくなった。分散分析の結果からは被験者 A ~ C については、壁の有無による有意差は認められなかったが、被験者 D ~ G については、壁の有無による有意差が認められた。

このように床面のような平面を一緒に描画し絵画的要因と組み合わせることで両眼での利用も多くのユーザに対して実現可能であることが示唆された。

表 5 両眼利用時の正規化誤差 (各被験者)

被験者		A	B	C	D	E	F	G
Near	壁なし	0.05	-0.10	-0.11	-0.53	-0.22	-0.33	-
	壁あり	-0.10	-0.05	-0.01	-0.33	-0.16	-0.20	-0.30
Middle	壁なし	-0.10	-0.38	-0.19	-0.55	-0.41	-0.48	-
	壁あり	-0.06	-0.31	-0.05	-0.43	-0.30	-0.40	-0.38
Far	壁なし	-0.24	-0.48	-0.27	-0.57	-0.45	-0.39	-
	壁あり	-0.26	-0.43	-0.20	-0.50	-0.38	-0.40	-0.44

表 6 2 要因分散分析結果 (各被験者)

被験者		A	B	C	D	E	F	G
片眼両眼	Near	両眼 **	片眼 <i>n.s.</i>	両眼 **	片眼 *	片眼 <i>n.s.</i>	片眼 <i>n.s.</i>	片眼 **
	Middle	両眼 **	片眼 <i>n.s.</i>	両眼 <i>n.s.</i>	片眼 <i>n.s.</i>	片眼 +	両眼 <i>n.s.</i>	片眼 <i>n.s.</i>
	Far	両眼 **	片眼 <i>n.s.</i>	両眼 *	片眼 <i>n.s.</i>	片眼 <i>n.s.</i>	両眼 *	片眼 <i>n.s.</i>
壁の有無	Near	無し <i>n.s.</i>	有り <i>n.s.</i>	有り <i>n.s.</i>	有り **	有り **	有り *	有り **
	Middle	有り <i>n.s.</i>	有り <i>n.s.</i>	無し <i>n.s.</i>	有り **	有り **	有り *	有り **
	Far	有り <i>n.s.</i>	有り <i>n.s.</i>	有り <i>n.s.</i>	有り **	有り **	有り <i>n.s.</i>	有り **
相互作用	Near	*	- <i>n.s.</i>	**	- <i>n.s.</i>	*	- <i>n.s.</i>	-
	Middle	- <i>n.s.</i>	- <i>n.s.</i>	**	- <i>n.s.</i>	*	- <i>n.s.</i>	-
	Far	- <i>n.s.</i>	- <i>n.s.</i>	*	- <i>n.s.</i>	*	- <i>n.s.</i>	-

n.s.:有意ではない +:有意傾向 $p < .1$ *:有意 $p < .05$ **:有意 $p < .01$

5.3 イメージベースシステムによる物体配置の知覚精度評価実験

イメージベースシステムにより運動視差の提示を行う場合，オンラインシステムでは遅れが生じてしまうため，遅れの問題を切り分けることが困難である．そこで遅れなく映像提示の可能な多視点画像データベースを用いたシステムを利用して遅延と立体感の影響について評価実験を行う．ただし多視点画像データベースのシステムでは離散的な画像を提示することになるため，提示画像の離散化による影響も検討する必要がある．始めに画像の提示間隔に対する立体感の評価を行い，離散的な画像提示による影響を評価する．次に擬似的に遅延を発生させ，立体感に対する映像遅延の影響について評価を行う．

5.3.1 多視点実画像の蓄積による運動視差提示実験方法

事前に撮影した画像を用いたイメージベースによる運動視差 3次元ディスプレイシステムを使用して，頭部移動に対して提示する画像の提示間隔，提示遅れに対する利用実験を行う．

実験システムの外観を図 28 に示す．アルミフレームの台に透明アクリル版が取り付けられた実験用ステージであり，下からキャリブレーションされたカメラを用いて物体の位置を計測することが出来る．実験用画像撮影時のオブジェクトの位置を事前に下からのカメラで撮影しておき，実験後のオブジェクトの位置と比較する．配置してもらう物体の底面には，それぞれ異なったマーカが貼られており，マーカの取り付けられた物体の位置はテンプレートマッチングにより追跡する．実験に用いたオブジェクトを図 29 に示す．運動視差を用いずに，円柱のサイズの変化から奥行感を得ることがないように，ユーザに見てもらう物体には円柱のオブジェクトを，ユーザに配置してもらう物体には板状の物体を利用した．物体の位置移動の記録は図 30 のインタフェースを用いて計測され，目標位置と現在位置が記録される．

大学院生男女 9 名の被験者に 3次元ディスプレイを通して 2つの円柱を見てもらい，同様の配置となるよう板状の物体を配置してもらい，立体感と配置のしやすさの 2点についてアンケートにより 4段階評価で数値をつけたもらった．実験

に用いた画像は、左右に 2mm 刻みで $\pm 150\text{mm}$ の 151 枚であり、この時のモニタまでの距離は 500mm である。実験で使った画像を図 31 に示す。評価は、配置された物体の位置の比較による客観評価とアンケートによる主観評価により議論する。



(a) 実験用画像の撮影



(b) 実験環境

図 28 イメージベースシステムによる物体配置の知覚評価実験システム

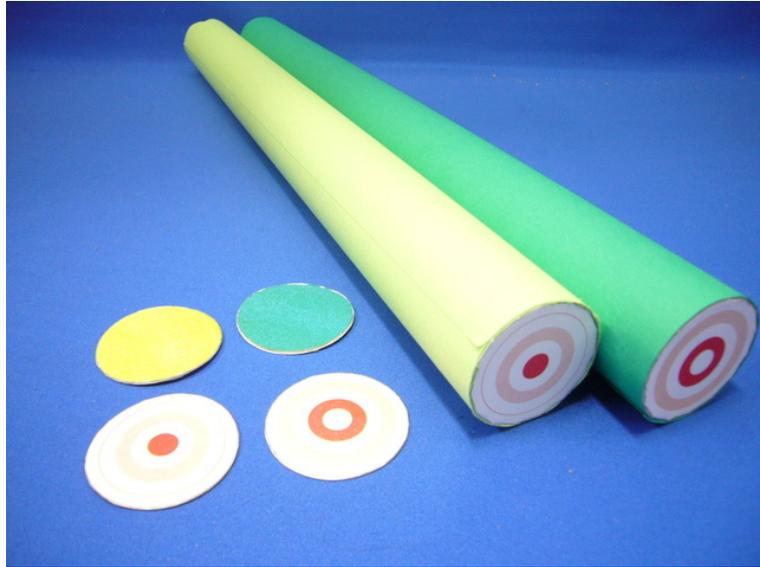


図 29 マーカ付き実験用オブジェクトと配置物体

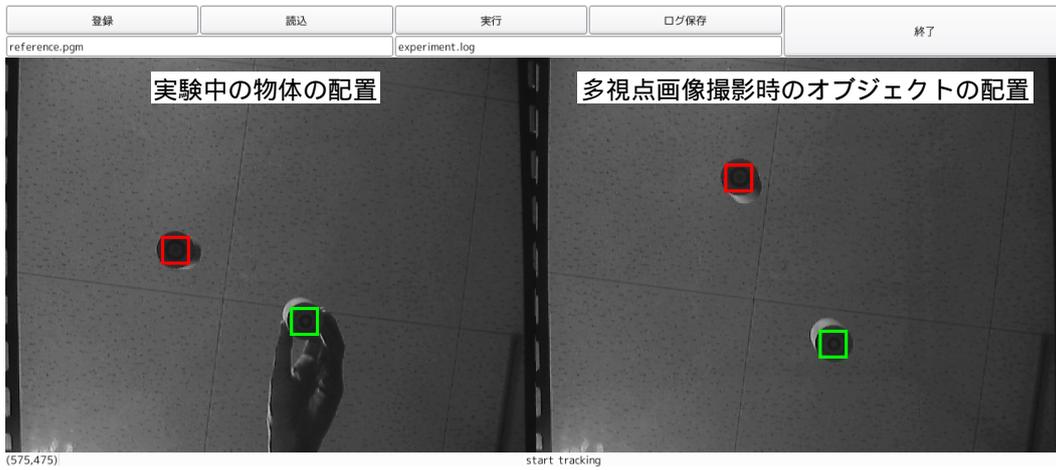
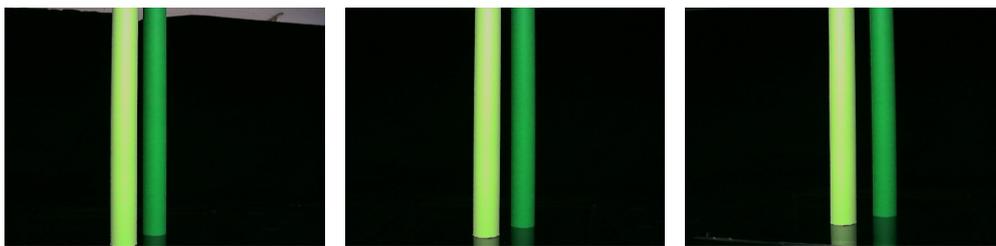


図 30 実験用オブジェクトの位置誤差計測インターフェース



(a) 左側からの画像

(b) 中央からの画像

(c) 右側からの画像

図 31 実験で使った画像の例

5.3.2 画像の提示間隔の評価

画像提示間隔を 150mm, 50mm, 10mm, 2mm と順番に変更し, 映像提示を行った。この時に提示された画像枚数はそれぞれ, 3 枚, 7 枚, 31 枚, 151 枚である。配置された 2 物体の左右距離, 前後距離の平均と標準偏差の結果を図 32 に示す。棒グラフは平均を, 誤差棒は標準偏差を, 横線はユーザに提示した 2 物体間の実際の距離である。

物体配置の結果について分散分析を行ったところ有意差は得られなかった。物体配置の結果を見ると, 左右方向の距離についてはほぼ正確な距離として知覚されているが, 前後方向の距離については近くに感じているという結果になった。Hayashibe らによれば VR 空間に限らず実画像による運動視差提示においても知覚距離は過小評価されるという報告 [41] があり, これが原因となり距離が近くに感じられたと考えられる。また画像の切替え幅が大きくなると知覚される距離も遠くなり, 標準偏差も大きくなっている。図 34 に画像の提示間隔が 150mm, 50mm の時の物体が存在するように感じる範囲を示す。画像の提示間隔が小さくなると物体の存在が感じられる範囲も狭くなり, 決まった位置に定位するようになる。そのため画像切替え幅が大きいときには標準偏差が大きくなると考えられる。

アンケートの結果の平均と標準偏差を図 33 に示す。アンケート結果について分散分析を行ったところ, 立体感, 配置しやすさ共に有意水準 1% で有意差が認められた。ライオン法 [45] を用いて有意水準 5% による多重比較を行った結果を表 7, 表 8 に示す。これ以外の組合せでは有意差は認められなかった。立体感については 2mm, 10mm それぞれについて 50mm, 150mm より立体感が高いという結果が得られたが, 配置しやすさについては, 2mm と 150mm の比較でのみ差が得られた。つまり画像の提示間隔が大きい場合でも物体を置くという操作を行う上ではそれほど支障はないが, 立体的には感じられていないことが伺える。物体配置の結果において有意差が得られなかったことはこの立体感と操作しやすさの違いによる影響も考えられる。

これらの結果から画像提示間隔が小さいほど立体感や操作しやすさにおいて良い結果を得ることが分かった。また物体を置くという操作を行うことと立体的に感じることに間に差が存在し, 立体に感じることの方がより厳しい条件であるこ

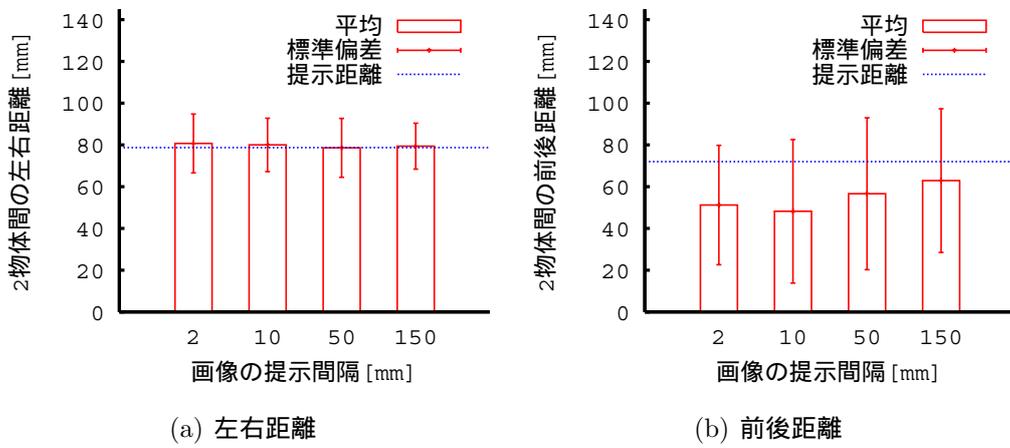


図 32 画像提示間隔の物体配置評価

とが確認された。

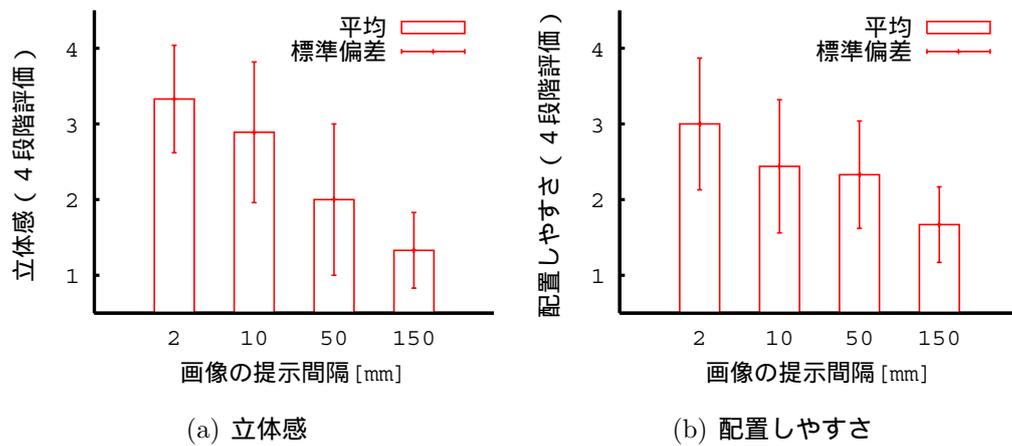


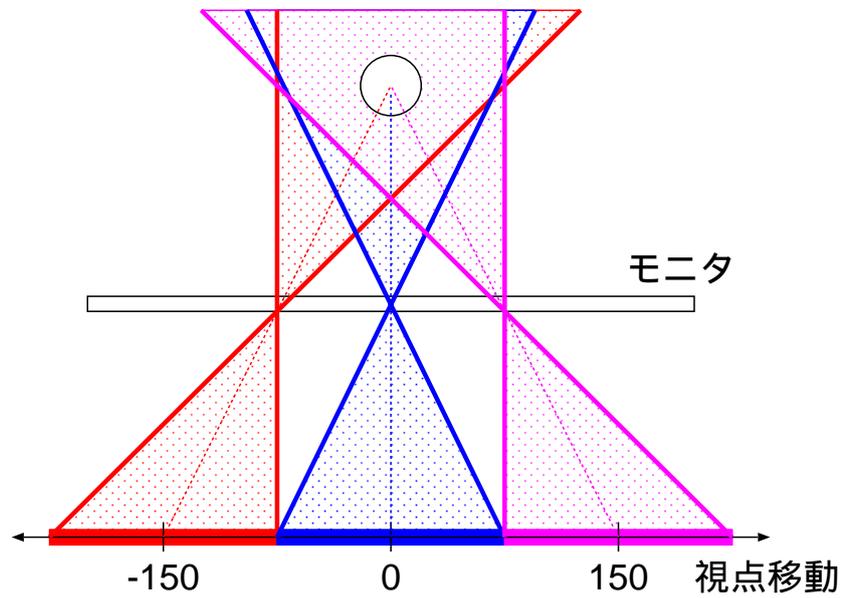
図 33 画像提示間隔のアンケートによる主観評価

表 7 立体感のライアン法による多重比較の結果

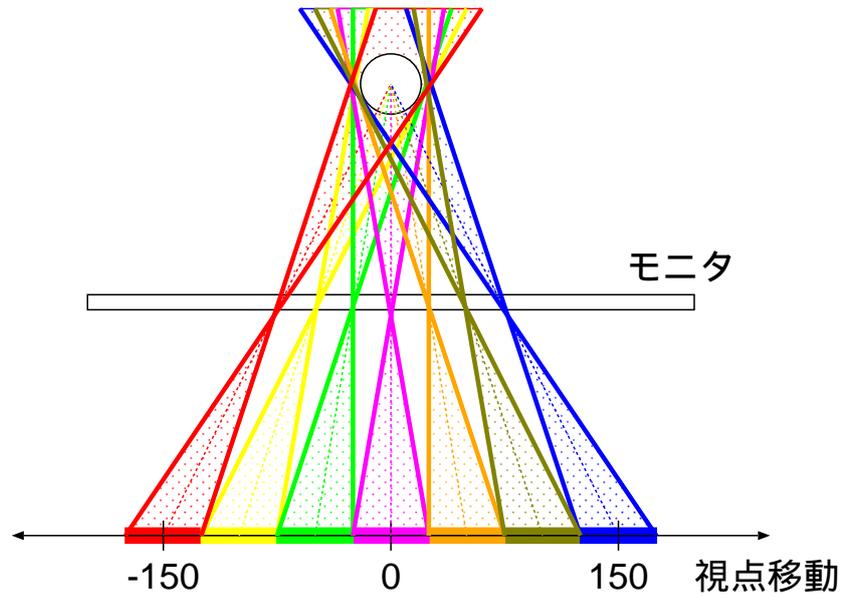
比較する群	名義水準	t 検定による有意確率	判定
150 2	0.0083	0.0001	有意
150 10	0.0125	0.0033	有意
50 2	0.0125	0.0039	有意
50 10	0.0250	0.0092	有意

表 8 配置しやすさのライアン法による多重比較の結果

比較する群	名義水準	t 検定による有意確率	判定
150 2	0.0083	0.0017	有意



(a)150mm間隔での運動視差提示



(b)50mm間隔での運動視差提示

図 34 画像提示間隔に対する物体知覚範囲

5.3.3 提示遅れによる知覚精度の評価

画像の提示間隔は小さい程，立体感が大きい結果が得られるため，画像提示間隔 2mm に対して，映像の提示遅れを 1000ms，500ms，250ms，125ms，遅れなしと順番に変更し，映像提示を行った．配置された 2 物体の左右距離，前後距離の平均と標準偏差の結果を図 35 に棒グラフは平均を，誤差棒は標準偏差を，横線はユーザに提示した 2 物体間の実際の距離である．

物体配置の結果について分散分析を行ったところ有意差は得られなかった．頭部動きが大きいときには遅れの影響が顕著に表れるが，物体配置を行うために頭部の動きが比較的ゆっくりになってしまい，あまり影響がでなかったものと考えられる．

アンケートの結果の平均と標準偏差を図 36 に示す．アンケート結果について分散分析を行ったところ，立体感，配置しやすさ共に有意水準 5% で有意差が認められた．ライオン法を用いて有意水準 5% による多重比較を行ったところ，立体感については図 9 のように 1,000ms と 0ms，500ms と 0ms，250ms と 0ms の組み合わせで有意差が得られ，他の組み合わせでは有意差は得られなかった．また配置しやすさについても有意差は得られなかった．画像の切替え幅の時と同様に画像の提示遅れについても立体感と配置しやすさについては差が存在し，立体感を感じるにはより厳しい条件であることが分かる．今回の結果では約 250ms での遅れに対してもユーザは違いを感じていることが確認された．遠隔操作などを行う場合，遅延は 500ms 以上になると操作性が悪くなるといわれている．この数値を比較しても立体感を得るにはより遅れの少ない映像提示が必要であると考えられる．現在構築したシステムは平均遅れが約 728ms 存在し，立体感をユーザに感じさせるには不十分であると考えられる．遅延が生じている原因はマニピュレータの動作遅れにあり，立体感の高い運動視差提示を行うためには，これを改善する必要があるだろう．

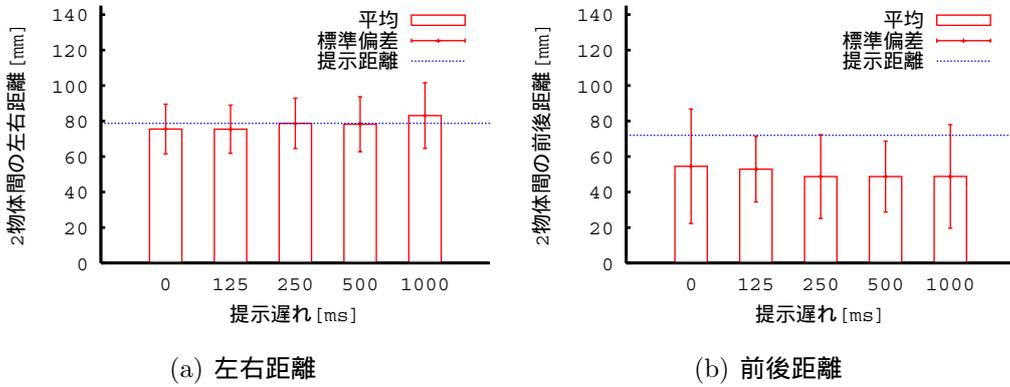


図 35 画像提示遅延の物体配置評価

表 9 立体感のライアン法による多重比較の結果

比較する群	名義水準	t 検定による有意確率	判定
1000 0	0.0050	0.0035	有意
500 0	0.0067	0.0011	有意
250 0	0.0100	0.0040	有意

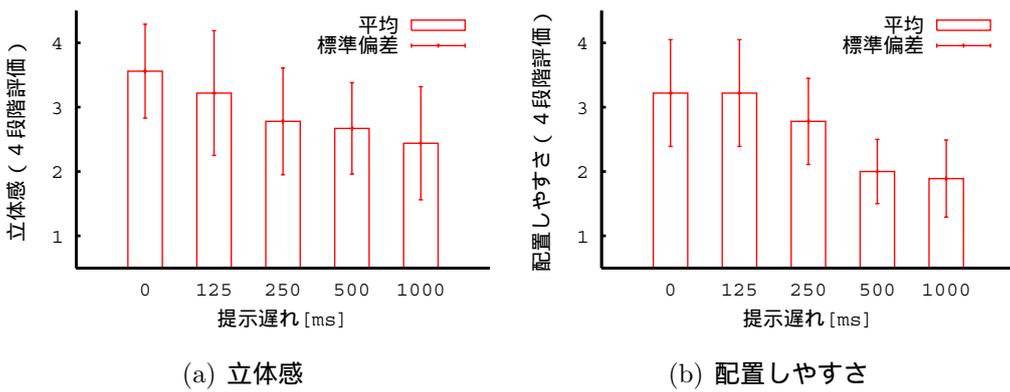


図 36 画像提示遅延のアンケートによる主観評価

5.4 本章の結論

構築したモデルベースによる運動視差 3次元ディスプレイシステムを利用した片眼・両眼における絶対距離知覚実験を行った。実験から本システムを利用して距離の違いを知覚できていることが確認された。また観察距離が大きくなるにつれて知覚距離の誤差が大きくなる症状が見られ、絶対距離知覚を必要とするような精密な作業には不向きであることが示唆された。また床面や壁などの平面物体と一緒に描画し絵画的要因と組み合わせた場合においては、両眼視差が立体知覚を邪魔してしまう両眼での利用時においても多くのユーザに対して定位置に知覚されることが示された。

イメージベースの実験では、画像の提示間隔と映像提示の遅延について評価を行い、画像の提示間隔、遅延ともに小さい程より立体感、操作感が向上することを示した。画像の提示間隔については、10mm から 50mm を超えると統計的に立体感が低下し、また遅れについては 250ms を超えると統計的に立体感が低下することを実験から示した。また実験結果から立体感と操作感には差があり、立体感を与える方がより難しいことを確認した。

6. 結論

6.1 まとめ

様々なシステムにおいて、より直感的なインタフェースが求められているなか、視覚デバイスにおいても様々な3次元ディスプレイが開発されてきた。しかし現状ではまだ高価であり手軽に利用できるものではなく、日常的に広く利用されているテレビやコンピュータのモニタについては2次元的にしか利用されていないのが現状である。

そこで本論文では、通常の平面モニタを利用して実現可能な3次元ディスプレイシステムとして視点位置計測に基づく運動視差3次元ディスプレイの提案を行い、CGモデルを用いるモデルベースシステムと多視点画像データベースを用いるシステム、カメラをオンラインで動作させるシステムの3つのシステムの実装を行った。モデルベースのシステム、多視点画像データベースのシステムでは約30msの遅れでの立体提示が可能である。またオンラインでのシステムは約728msという映像遅れが存在するが、動物体に対する立体提示が可能である。これらのシステムは用途による使い分けることが適切である。例えば、モデルルームなどのCADモデルに基づき設計された部屋をCGにより閲覧する場合などはモデルベースのシステムが利用可能であり、ブラウザを用いての3Dビューの提示を行う場合などには多視点画像データベースのシステムの利用が想定される。実世界の動物体を閲覧する場合などには、オンラインでのシステムの利用が想定される。

運動視差を用いた奥行き知覚実験やユーザアンケートにより、本システムを利用して奥行感が得られることを確認した。ただし遠くの物体ほど物体までの距離の誤差が大きくなるため、絶対距離知覚を必要とするような精密な作業には不向きであることが示唆された。また画像データベースを用いるような離散的な画像を提示する場合には、画像の提示間隔が10mmから50mmを超えると統計的に立体感が低下することを示し、オンラインシステムのように映像の提示遅れが大きくなるシステムでは250msを超えると立体感が低下することを示した。

6.2 今後の課題と展望

モデルベースのシステムでは提示物体の色や照明，陰影については任意の条件を用いており，実環境とは異なった条件となっている．通常，色には進出色や後退色などがあり同じ場所の物体を見たときでも，赤や黄色は前方に，青や黒などは後方に，緑色は中央に知覚されるなど色により距離感が異なる．また照明や陰影についても同様に距離感に影響を及ぼすと言われている [3][5]．これらを考慮し，より現実に近いCGの描画を行うことで，奥行き感の向上が見込まれる．

またイメージベースのシステムでは，画像間の視点における中間画像の生成は行っておらず，必要な画像は全て取得しておくかカメラを動かすことで運動視差提示の実現を行っている．そのため，事前に取得していない視点の画像は提示することが出来ず，またカメラを動作させた場合にも動作遅れにより映像提示にも遅延が生じてしまっている．モーフィングや光線空間法を用いることで中間画像の生成を行うことが可能であり，数台のカメラとアクティブカメラのハイブリットなシステムにした場合，複数台カメラに囲まれた視点範囲での自由視点をカメラ移動遅れの影響を受けずに提示することができ，またカメラ台数をあまり増やさなくとも様々な視点での映像提示が可能になると考える．こういった遅れが少なくなかつ広範囲での任意視点における運動視差の提示は今後の課題である．

本研究で実装したオンラインシステムでは，画像の伝送を行っておらず車やロボットなどの運動席から周囲の映像を提示するといったカメラと映像提示が一体化された場面でのみ利用が可能である．遠隔モニタリングのようなアプリケーションのための映像伝送については今後の課題である．映像の伝送方法についてはH.323プロトコルなどの標準化などが行われており，近年のビデオ会議システムなどで利用されている．また，より臨場感のある立体提示を行うためには高解像度であるのが望ましいが，映像伝送の観点では遅延低減などの理由から伝送量の削減が必要となり，相反する二つの要望を共に満たすことは困難である．伝送量の削減方法としては，画像中心部や注目点付近を高解像度で伝送し，周辺は解像度を低下させることで画質低下を抑えながら伝送量の削減を行っている研究 [46][47][48] や複数視点画像に対して空間領域，時間領域での冗長性に着目したハイブリットなDCT(離散コサイン変換)符号化 [49] などがある．現在はまだ使用

していないが、視点位置計測として利用した顔情報計測システムでは視線方向の計測も可能である。このような視線情報を利用して画像内での注視情報を推定し、注視点付近を高解像度に、周辺視野部を低解像度に提示を行うことで、伝送量の削減などを行うことも今後行っていきたい。

2008年1月7日、総務省は2025年の家庭向け立体放送開始を目指し、立体ディスプレイの開発を民間企業と共同で行うことを発表した。今後も立体映像の提示技術のみならず、立体映像生成技術や効率的な立体映像の伝送技術など多くの研究開発、製品化が行われることであろう。本研究の成果がこの立体ディスプレイ開発の今後の発展に影響を与えることができるのであれば、筆者にとって大きな喜びである。

謝辞

本論文に書き記した一連の研究は、筆者が奈良先端科学技術大学院大学 情報科学研究科在学中に、ロボティクス講座（小笠原研究室）において執り行われたものです。この研究は多くの方々のお力添えにより実現することができました。ここでお礼を申し上げたいと思います。

本学 情報科学研究科 小笠原 司 教授は、修士、博士課程の5年間を通して、筆者の研究を温かく見守っていただくとともに、研究者としてのあり方を未熟な筆者に対して辛抱強くご指導頂きました。また研究活動のみならず、知識の足りない修士1年の頃から計算機の管理をしたいという筆者の希望を聞き入れていただき、情報技術に関わる広い知識を付ける機会を与えてくださいましたこと、合わせて深く感謝いたします。

本学 情報科学研究科 松本 健一 教授には、本研究をまとめるにあたり、数々のご助言、ご指摘を頂きました。多くの課題の中で実際に手を動かすべき課題の切り分けが出来ましたのも、松本先生のお力添えのおかげです。ありがとうございました。

本学 情報科学研究科 松本 吉央 准教授は、本研究の立ち上げから遂行に当たり、様々なご助言、ご指導を頂きました。投稿論文や英語論文執筆、発表練習の際には、昼夜問わず何度も添削をしていただきましたこと大変感謝しています。また愛知万博をはじめとする様々なイベントの参加や企業との共同研究など大変貴重な経験をさせて頂きました。公私共に本当に御世話になりました。ありがとうございました。

本学 情報科学研究科 横矢 直和 教授は、本研究を遂行する上で大変貴重なご助言、ご指摘を頂き、本研究をまとめる上での不足点を明確にすることが出来ました。また急な委員のお願いも快く引き受けて頂き大変感謝しています。ありがとうございました。

本学 情報科学研究科 上田 淳 助教は、本研究を進める上で、数学的な話題や本研究に対する率直な疑問点など様々なご助言を頂きました。ありがとうございました。

奈良産業大学 上田 悦子 准教授は、本研究を進める上での技術的なご助言にと

どまらず、研究を遂行にあたって不安になった時など相談に乗っていただき精神的にも大変お世話になりました。ありがとうございました。

本学 情報科学研究科 栗田 雄一 助教は、本研究に対する利点、問題点に対して鋭いご助言を頂き、本研究の不足な点などを明確にすることが出来ました。また社会人としての厳しさや研究者としての視点など大変勉強させて頂きました。ありがとうございました。

秘書 金岡 恵 様には、物品の発注や出張、雇用に係わる書類など様々な事務処理の手続きのサポートをして頂きました。また2度に渡る NAIST 卓球大会出場や市民大会への参加の時には団体戦メンバーとして、ダブルspartnerとして多くの御支援を頂きました。ありがとうございました。

本学 情報科学研究科 怡土 順一 研究員、立命館大学 小枝 正直 講師は、プログラミングテクニックやサーバ構築に関する議論をはじめ、アルゴリズムに対する指摘など技術面において大変御世話になりました。情報技術に係わる多くの知識を楽しみながら身に付けることが出来たのもお二人のおかげです。ありがとうございました。

奈良工業高等専門学校 竹村 憲太郎 助教は、本研究の立ち上げから様々な御指導、御意見を頂き、投稿論文や博士論文の執筆の際にも適切な御助言を頂きました。また公私共に、大きな事から小さな事まで様々な事を気に掛けて頂き、多くのご支援を頂きましたこと大変感謝しています。ありがとうございました。

本学 情報科学研究科 2006 年度修了の 近藤 誠宏 様には、マニピュレータ制御のプログラムを作成する際に多くのご助言、ご支援を頂きました。ありがとうございました。

本学 情報科学研究科 2006 年度修了の Abdelaziz Khiat 様、博士後期課程 Albert Causo 君には、論文執筆や発表資料作成の際に英語文章の添削を手伝っていただきました。また同じ博士後期課程とし時には励まし合いながら研究活動を行い、大変お世話になりました。ありがとうございました。

本学 情報科学研究科 博士後期課程の 丁 明 君、池田 篤俊さんは、同じ博士後期課程としての研究を始め、デモやロボットを通じた地域交流などさまざまな研究室での活動を共に励まし合い、時には助け合いながら行いました。本当にお世

話になり、ありがとうございました。

本学 情報科学研究科 博士前期課程 玉井 康之 君には、実画像提示におけるマニピュレータの制御や多視点画像データベースを利用した実験などをして頂きました。多視点の実画像撮影や実験方法の設定では、うまくデータが出ず苦労することもあり多くの議論と時間を重ねましたが、何とかまとめることが出来ました。ありがとうございました。

本学 情報科学研究科 博士前期課程 倉木 大輔 君、小林 純也 君、松尾 真依さん、浅野 慧 君、山本 和樹 君、野村 尚央 君は同じ研究グループとして研究の本質から些細なことまで多くのことを語り合いました。たびたび議論が白熱し、ミーティングが長丁場となることもありましたが、それゆえにお互いの研究を深く理解し合い、より良い研究活動ができたと思います。ありがとうございました。

また研究室生活を送るにあたり、様々な形で御支援、御協力、御鞭撻を頂きました本学 ロボティクス講座の卒業生および在学生のすべての方々にお礼申し上げます。

最後に、5年間にも渡る大学院生活を送るにあたり協力、応援してくれた家族に深く感謝します。

参考文献

- [1] 池内克史, 倉爪亮, 西野恒, 佐川立昌, 大石岳史, 高瀬裕: The Great Buddha Project -大規模文化遺産のデジタルコンテンツ化; 日本バーチャルリアリティ学会論文誌, vol.7, no.1, pp.103-123 (2002).
- [2] 坂村健: デジタルミュージアムからユビキタスミュージアムへ (<特集> デジタルアーカイブ); 人工知能学会誌, vol.18, no.3, pp.259-266 (2003).
- [3] 大山正: 視覚心理学への招待 -見えの世界へのアプローチ-, 第7章, pp.147-181, サイエンス社 (2000).
- [4] 竹市博臣: 知覚と運動, 第1章, pp.15-48, 東京大学出版会 (1995).
- [5] 高木幹雄: 新編画像解析ハンドブック, 第II-1章, pp.391-472, 東京大学出版会 (2004).
- [6] Doug A. Bowman, Ernst Kruijff, Jr Joseph J. LaViola, and Ivan Poupyrev: 3D ユーザインタフェース, 第3章, pp.32-64, 丸善株式会社 (2005).
- [7] 増谷健, 東野政弘, 安東孝久, 金山秀行: 解説: メガネなし 3d ディスプレイの開発; システム制御情報学会誌, vol.49, no.9, pp.351-356 (2005).
- [8] Ken Perlin, Salvatore Paxia, and Joel S. Kollin: An Autostereoscopic Display; Proceeding of SIGGRAPH 2000 (2000).
- [9] Glenn Bresnahan, Raymond Gasser, Augustinas Abaravichyus, Erik Brisson, and Michael Waltherman: Building a Large Scale, High Resolution, Tiled, Rear Projected, Passive Stereo Display System Based on Commodity Components; Proceedings of SPIE, vol.5006, pp.19-30 (2003).
- [10] 佐藤誠, 長谷川晶一, 橋本直己: マルチモーダル仮想環境のためのマルチプロジェクションシステム D-vision の開発; 画像電子学会第197回研究会 高臨場感ディスプレイフォーラム 2002 (2002).

- [11] 柴野伸之, 柏木正徳, 澤田一哉: 半球ドーム型 VR システム「CyberDome」; Technical report, 松下電工技報 (2003).
- [12] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti: Surround-screen projection-based virtual reality: The design and implementation of the CAVE; ACM SIGGRAPH'93 Proceedings, pp.135-142 (1993).
- [13] Michitaka Hirose, Tetsuro Ogi, Shohei Ishiwata, and Toshio Yamada: Development and Evaluation of Immersive Multiscreen Display "CABIN"; Systems and Computers in Japan, Scripta Technica, vol.30, no.1, pp.13-22 (1999).
- [14] 山田俊郎, 棚橋英樹, 小木哲朗, 廣瀬通孝: 完全没入型 6 面ディスプレイ COSMOS の開発と空間ナビゲーションにおける効果; 日本バーチャルリアリティ学会論文誌, vol.4, no.3, pp.531-538 (1999).
- [15] Kenji Tanaka, Junya Hayashi, Yutaka Kunita, Masahiko Inami, Taro Maeda, and Susumu Tachi: "TWISTER" Technical Challenges; ACM SIGGRAPH2002 Conference Abstract and Applications, p.271 (2002).
- [16] M. Agrawala, A. C. Beers, B. Frohlich, and P. Hanrahan: The two-user responsive workbench : support for collaboration through individual views of a shared space; Computer Graphics, Annual Conference Series (Proc. of SIGGRAPH), pp.327-332 (1997).
- [17] 北村喜文, 小西孝重, 山本澄彦, 岸野文郎: 人数共有型立体ディスプレイ IllusionHole; 映像情報メディア学会誌, vol.57, no.10, pp.1320-1327 (2003).
- [18] 島田悟, 木村龍実, 欠端雅之, 佐々木史雄: 空中に浮かび上がる 3次元 (3D) 映像; 産総研 TODAY, 第 6 巻, pp.16-19, , 産業技術総合研究所 (2006).
- [19] Levent Onural, Thomas Sikora, Jorn Ostermann, Alijoscha Smolic, M. Reha Civanlar, and John Watson: An Assessment of 3DTV Technologies; NAB Broadcast Engineering Conference Proceedings 2006, pp.456-467 (2006).

- [20] James E. Cutting, and Peter M. Vishton: Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth; Perception of Space and Motion(W. Epstein and S. Rogers, eds.), pp.69-117, Academic Press (1995).
- [21] 川上直樹, 稲見昌彦, 柳田康幸, 前田太郎, 舘暲: オブジェクト指向型ディスプレイの研究; vol.40, no.6, pp.2725-2733 (1999).
- [22] Jurriaan D. Mulder, and Breght R. Boschker: A Modular System for Collaborative Desktop VR/AR with a Shared Workspace; Proceedings of the IEEE Virtual Reality Conference 2004, pp.75-82 (2004).
- [23] Siegmund Pastoor, Jin Liu, and Sylvain Renault: An Experimental Multimedia System Allowing 3-D Visualization and Eye-Controlled Interaction Without User-Worn Devices; IEEE Transactions on Multimedia, vol.1, no.1, pp.41-52 (1999).
- [24] 河原塚有希彦, 高橋誠史, 宮田一乗: ViewFrame2 -マーカレス顔部検出手法を利用した”ViewFrame”-; 芸術科学会論文誌, vol.3, no.3, pp.189-192 (2004).
- [25] J. P. Djajadiningrat, C. J. Overbeeke, and P. J. Stappers: Cubby: A Unified Interaction Space for Precision Manipulation; Proceedings of ITEC2001 (CD-ROM) (2001).
- [26] 松本吉央, 怡土順一, 竹村憲太郎, 小笠原司: リアルタイム顔・視線計測システムの開発と知的インタフェースへの応用; 情報処理学会論文誌コンピュータビジョンとイメージメディア, vol.47, no.SIG 15(CVIM16), pp.10-21 (2006).
- [27] Takeo Kanade, Peter Rander, and P. J. Narayanan: Virtualized Reality: Constructing Virtual Worlds from Real Scenes; IEEE Multimedia, vol.4, no.1, pp.34-47 (1997).
- [28] P. J. Narayanan, Peter Rander, and Takeo Kanade: Constructing Virtual Worlds Using Dense Stereo; the International Conference on Computer

- Vision (ICCV) 98, pp.3-10 (1998).
- [29] Andrew Lippman: Movie-Maps: An Application of the Optical; ACM SIGGRAPH Computer Graphics, vol.14, no.3, pp.32-42 (1980).
- [30] Yasuyuki Yanagida, Shintaro Saito, Seiichiro Yano, Taro Maeda, and Susumu Tachi: A Head-Tracked, Live-Video-Based Telexistence System Using a Fixed Screen; Proceedings of ICAT 2001 (2001).
- [31] 財津義貴, 中川高志, 稲見昌彦, 川上直樹, 柳田康幸, 前田太郎, 舘暲: テレグジスタンスの研究(第32報)-頭部運動に高速追従するトルソ型撮像ロボットの開発-; 信学技報 MVE2001-11, pp.51-54 (2001).
- [32] Steven M. Seitz, and Charles R. Dyer: View Morphing; Proc. SIGGRAPH 96, pp.21 - 30 (1996).
- [33] 石川智也, 山澤一誠, 横矢直和: 複数の全方位動画像からの自由視点画像生成; MIRU2005, pp.1396-1403 (2005).
- [34] 苗村健, 金子正秀, 原島博: 光線情報の正投影表現に基づく3次元空間の記述; テレビジョン学会誌, vol.51, no.12, pp.2082-2090 (1997).
- [35] 福嶋慶繁, 圓藤知博, 藤井俊彰, 谷本正幸: 光線空間法に基づく実時間広角自由視点画像合成; 信学技法 SIP2005-84, pp.61-66 (2005).
- [36] 石川寛享, 北原格, 大田友一: 大規模空間の多視点映像を用いた運動視差の再現可能な自由視点映像提示; 信学技法 PRMU2000-190, pp.31-38 (2001).
- [37] 飯尾淳, 谷田部智之, 比屋根一雄: 顔の位置と頭部位置の3次元情報を利用したユーザインタフェースの試作; 第8回画像センシングシンポジウム SSII2002 (2002).
- [38] Shinjiro Kawato, Nobuji Tetsutani, and Kenichi Hosaka: Scale-Adaptive Face Detection and Tracking in Real Time with SSR Filters and Support

- Vector Machine; IEICE Trans. on Info. & Sys., vol.E88-D, no.12, pp.2857-2863 (2005).
- [39] 大塚聡子: 運動視差による奥行知覚の研究 Ph.D thesis , 東京大学文学部・大学院人文社会系研究科 (1998).
- [40] 西田委甲子 , 五十嵐智 , 阿刀田央一: 運動視差による立体知覚と定位の実験; 電子情報通信学会論文誌 D , vol.J71-D , no.4 , pp.740-742 (1988).
- [41] Keikichi Hayashibe: Apparent distance in actual, three-dimensional video-recorded, and virtual reality; Perceptual and Motor skills, vol.95, no.2, pp.573-582 (2002).
- [42] Jeanette C. Meng, and H. A. Sedgwick: Distance perception mediated through nested contact relations among surfaces; Perception and Psychophysics, vol.63, no.1, pp.1-15 (2001).
- [43] T. S. Meese, and D. J. Holmes: Performance data indicate summation for pictorial depth-cues in slanted surfaces; Spatial Vision, vol.17, no.1-2, pp.127-151 (2004).
- [44] George Mather, and David R. Smith: Combining depth cues: effects upon accuracy and speed of performance in a depth-ordering task; Vision Research, vol.44, pp.557-562 (2004).
- [45] 森敏昭 , 吉田寿夫: 心理学のためのデータ解析テクニカルブック , 第 4 章 , pp.171-172, 北大路書房 (1990).
- [46] 藤野雄一 , 橋爪誠 , 田上和夫 , 小西晃造 , 家入里志 , 清水周次 , 中島直樹 , 岡村耕二 , 上田幸宏 , 下野美智留 , 光石衛 , 割澤伸一 , 荒田純平 , 高橋弘樹 , Phongsan Pitakwatchara: グローバルメディカルテレインストラクション構想と日本-タイ間スケラブル codec ベース遠隔手術実証実験; 電子情報通信学会通信ソサイエティ大会 , pp.S-35,S-36 (2006).

- [47] 家永貴史, 松永勝也, 志堂寺和則, 合志和晃, 松木裕二, 永田裕樹: 中心部に高解像度部分を持つ立体ビデオシステムにおける映像伝送チャンネル削減の試み; 日本バーチャルリアリティ学会論文誌, vol.7, no.1, pp.15-21 (2002).
- [48] 鷹居頼治, 浜本隆之, 半谷精一郎: 人間の注視特性を考慮した低ビットレート画像符号化; 電子情報通信学会情報・システムソサイエティ大会, p.123 (1997).
- [49] 岡慎也, ナバンチャンプリム, 藤井俊彰, 谷本正幸: 自由視点テレビのための光線空間情報圧縮; 信学技報 CS2003-141, pp.7-12 (2003).
- [50] 吉川恒夫: ロボット制御基礎論, 第2章, pp.15-19, コロナ社 (1988).

著者研究業績

論文誌

1. 末永 剛, 松本 吉央, 小笠原 司: “非拘束な運動視差提示 3次元ディスプレイの提案と評価”, ヒューマンインタフェース学会論文誌, Vol.9, No.1, pp.49-56, 2007.2. (3-5章に対応)

国際会議 (査読あり)

1. Tsuyoshi Suenaga, Yasuyuki Tamai, Yuichi Kurita, Yoshio Matsumoto, and Tsukasa Ogasawara : “Poster: Image-based 3D Display with Motion Parallax using Face Tracking”, Proceedings of 3DUI2008, 2008.3. (3,4章に対応)
2. Tsuyoshi Suenaga, Yoshio Matsumoto, and Tsukasa Ogasawara : “Development and Evaluation of a Non-Contact 3D Display Based on Motion Parallax”, Proceedings of 12th International Conference on Human-Computer-Interaction international 2007 (HCII2007), pp494-498, 2007.7. (4,5章に対応)
3. Tsuyoshi Suenaga, Yoshio Matsumoto, and Tsukasa Ogasawara : “3D Display Based on Motion Parallax Using Non-contact 3D Measurement of Head Position”, In Proceedings of OZCHI2005, 2005.11. (3,4章に対応)

全国大会

1. 末永 剛, 松本 吉央, 小笠原 司 : “運動視差提示 3次元ディスプレイ -絶対距離知覚による評価実験-”, ヒューマンインタフェースシンポジウム 2006, 2242, 2006.9. (4,5章に対応)

2. 末永 剛, 竹村 憲太郎, 松本 吉央, 小笠原 司 : “視点位置計測を用いた半透明3次元ディスプレイ”, ヒューマンインタフェースシンポジウム2004, pp.939-942, 2004.10. (3,4章に対応)
3. 玉井 康之 末永 剛 栗田 雄一 松本 吉央 小笠原 司 : “運動視差提示による実画像3次元ディスプレイ”, 第25回日本ロボット学会学術講演会予稿集, 1H12, 2007.9.13-15. (3,4章に対応)

受賞

1. ヒューマンインタフェース学会 ヒューマンインタフェースシンポジウム2006 学術奨励賞 . “運動視差提示3次元ディスプレイ - 絶対距離知覚による評価実験 - ” . 2007年3月 . (4,5章に対応)

他, 国内発表 11 件

付録

A. 顔情報計測システム

図 37 に頭部位置視線計測システムのソフトウェア構成の概要を示す。

顔モデルの作成 顔のトラッキングを行う前に，個人毎に 3 次元顔モデルを作成する必要がある．3 次元顔モデルは，以下の 3 種類のデータから構成される．

- 特徴領域の 3 次元座標 (x, y, z)
- 特徴領域のテンプレート画像
- 顔全体の粗いテンプレート画像

まず，図 38 のような 1 枚の登録画像をキャプチャする．その画像からトラッキングに用いられる局所領域をマウス操作により指定し，登録する．特徴領域はトラッキングを行うために顔画像の中で十分に特徴的である領域を選択する．本研究では，ユーザの左右の目の両端，口の両端，眉の内側の端の合計 8 点を特徴領域として使用している．特徴領域のテンプレート画像を図 39 に示す．各特徴領域（右カメラ画像）は左カメラの全体画像に対し，正規化相関演算を行いステレオ画像の対応づけが行われ，三次元座標がそれぞれの特徴領域について求められる．また各特徴領域の三次元座標は顔モデルの値として保存され，顔モデルは 8 の点を持つ剛体として扱われる．

特徴領域のテンプレートは局所的な画像であるため，画像全体から顔領域を探索するには向いていない．そこで，図 38 のような登録画像から，図 40 のような顔全体の荒いテンプレート画像を作成する．顔全体のテンプレート画像のサイズは 60×30 [pixel] であり，処理時間を短縮するため低解像度となっている．また，照明変動の影響を押さえるために微分画像を用いる．

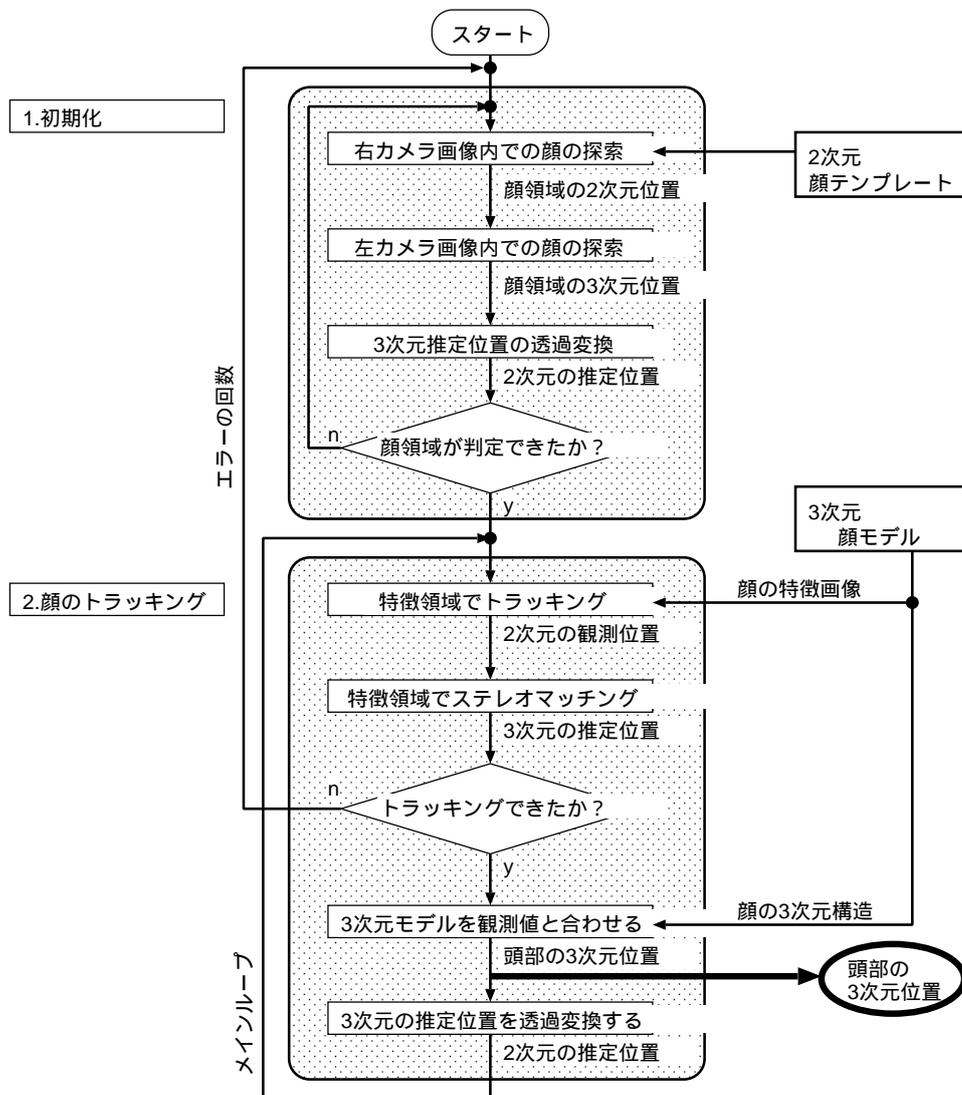


図 37 頭部位置視線計測システムのソフトウェア構成図



図 38 登録した顔画像

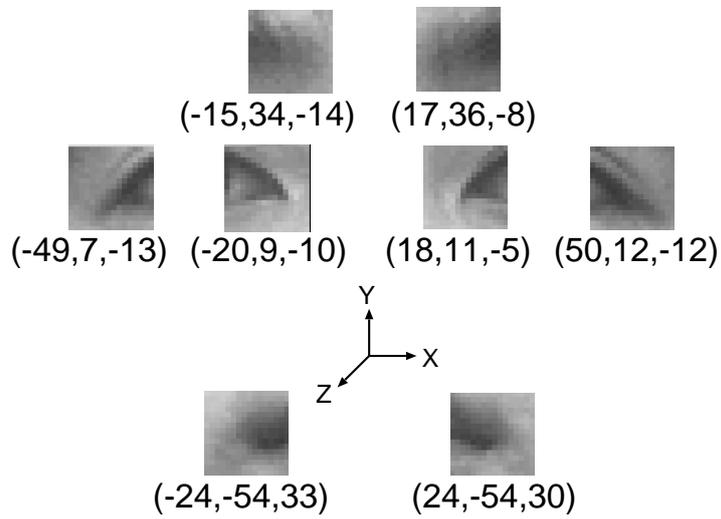


図 39 特徴領域のテンプレート画像



図 40 顔全体のテンプレート画像

初期化部 まず図 40 の顔全体のテンプレートを用いて、右カメラの画像中から顔の候補を探索する。候補が見つかった場合、右カメラ画像の中の顔候補領域をテンプレートとして、左カメラの画像で同様の探索を行う。これにより、顔領域の三次元位置が求められる。ただし、この三次元位置低解像度画像を用いているため、三次元位置の分解能も低い。そのため、この三次元位置はトラッキング開始前のシステムの初期化やトラッキング失敗からの復帰に用いられる。

トラッキング部 顔のトラッキングにおける三次元モデルマッチングは以下の手順で行われる。

1. 前フレームにおける計測姿勢をもとに、それぞれの特徴領域について、トラッキングにより 2 次元の観測位置を求める。
2. それぞれの特徴領域について、ステレオマッチングにより三次元の観測位置を求める。
3. 三次元モデルを三次元の観測値にフィッティングさせ、三次元の推定位置とする。
4. 三次元推定位置を透視変換して 2 次元の推定位置を求め、次のトラッキングでの探索範囲の基準とする。

モデルと観測値の重み付きフィッティングは、モデルと観測値の間に仮想的なバネを設定することで実現している。トラッキングとステレオマッチングの結果

から，各特徴領域につき相関値 (0 ~ 1) が 2 つ得られ，この相関値 2 つの積をそれぞれのバネの強さとする，トラッキングとステレオマッチングがどちらも良い場合には強いバネに，どちらも悪い場合には弱いバネになる．このバネ系が全体として最も小さなエネルギーを持つモデルの位置・姿勢を最適なフィッティング結果とすれば，モデルは高い相関が得られた観測値に強く引っ張られることになり，重み付けのかかったフィッティングが行われる．また，トラッキングとステレオマッチングのための正規化相関演算には MMX 命令を用いることで高速化をはかっている．

視線計測部 視線方向は，頭部の位置・姿勢と，眼の虹彩の位置との関係により決定される．その過程で用いられる三次元眼球モデルは，以下のパラメータからなる．

- 顔モデルに対する眼球中心の相対位置
- 眼球の半径
- 虹彩の半径

顔モデルに対する眼球中心の相対位置は，眼部分のトラッキングに用いられている眼両端の特徴領域の midpoint から眼球中心へ向かう三次元ベクトルとして与えられる．また，眼球の半径はおよそ 13[mm]，虹彩の半径は 7[mm] 程度と設定しており，現状では人手によって調整されている．

三次元の視線方向を計測する手順を以下に示す (図 41) ．

1. 眼球中心の三次元位置を，頭部の位置・姿勢の計測結果と，目の両端の特徴の midpoint から眼球中心に向かう三次元ベクトルを用いて決定する．
2. 眼球の三次元位置から，カメラ画像上での眼球の大きさが得られる．
3. 最も輝度値の和が小さい領域を求める事により虹彩の中心を算出する．
4. 画像上での眼球の中心と半径，および虹彩の中心の位置関係から視線の三次元ベクトルを求める．

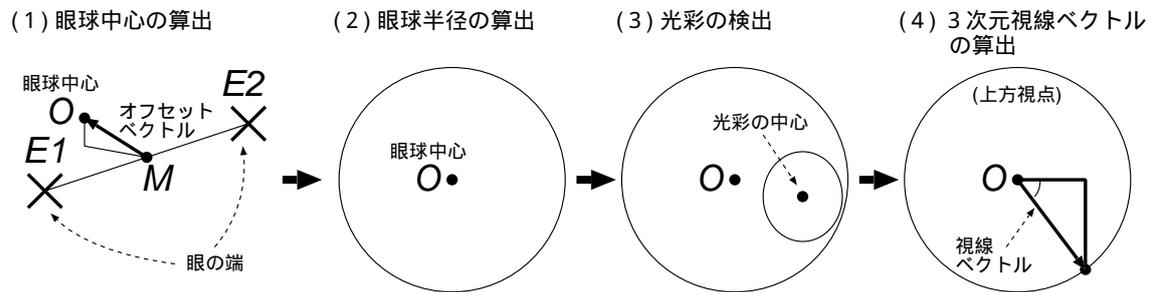


図 41 視線モデル

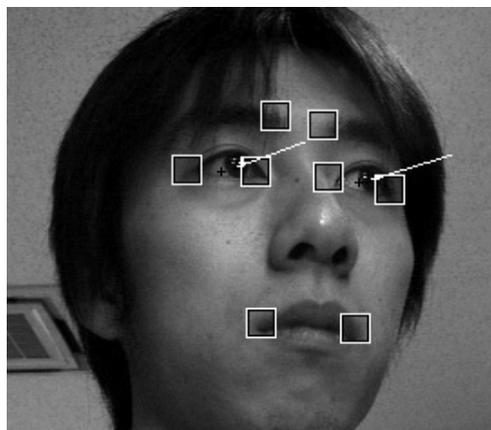


図 42 顔，視線計測結果

このようにして、頭部位置・姿勢と視線方向が計測され、次のフレームの始めには再び顔トラッキング部へ実行が移る。画像のキャプチャサイズは $320 \times 240[\text{pixels}]$ で、顔の位置の計測精度は約 $2[\text{mm}]$ 、顔の姿勢の計測精度は約 $2[\text{deg}]$ 、視線ベクトルの計測精度は約 $5[\text{deg}]$ 程度である。また、処理速度については、PC やカメラの性能にもよるが NTSC カメラを用いたトラッキングでは、ビデオレート ($30[\text{Hz}]$) を実現している。顔、視線計測結果を図 42 に示す。目から伸びている白い直線は視線ベクトルである。

B. 7軸マニピュレータPA10の仕様

表 10 マニピュレータ PA10 の仕様

型式	PA-10A-ARM		
形式	垂直多関節系		
関節数	7		
関節構成	ロボット取り付け面より R-P-R-P-R-P-R (R は回転関節, P は旋回関節)		
関節名称	ロボット取り付け面より S1-S2-S3-E1-E2-W1-W2 (S は肩関節, E は肘関節, W は手首関節を示す)		
アーム長	肩リーチ: 315mm (ベース面 ~ S2 間) 上腕: 450mm (S2 ~ E1 軸間) 下腕: 500mm (E1 ~ W1 軸間) 手首リーチ: 80mm (W1 ~ メカニカルインタフェース面間)		
関節動作範囲 及び 最高動作速度	軸名称	メカリミット (deg)	最高動作速度 (rad/sec)
	S1(回転)	±180	±1
	S2(旋回)	±94	±1
	S3(回転)	±180	±1
	E1(旋回)	±143	±2
	E2(回転)	±270	±2π
	W1(旋回)	±180	±2π
W2(回転)	±無限回転	±2π	
最大合成速度	1550 mm/sec		
可搬質量	10 kg		
駆動方式	AC サーボモータ		
センサ	出力軸 ブラシレスレゾルバ		

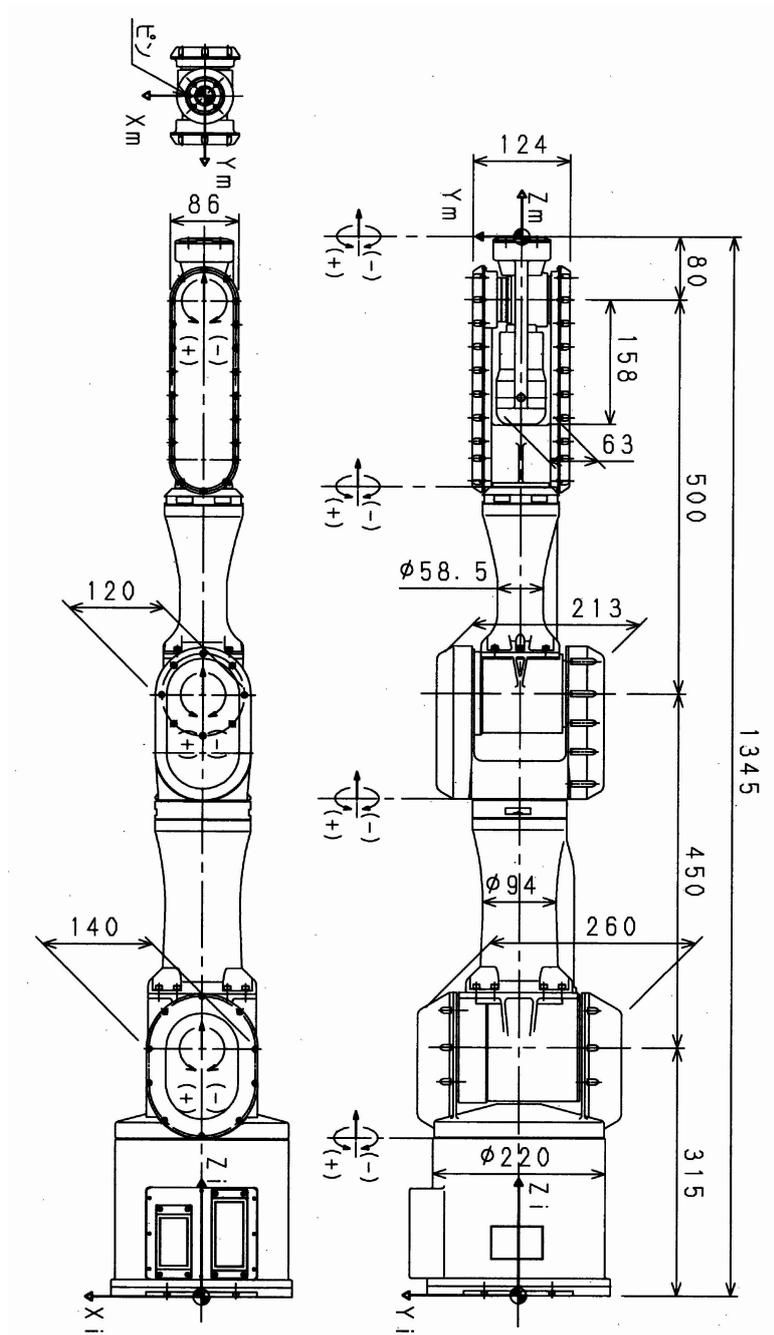


図 43 マニピュレータ PA10 の関節構造

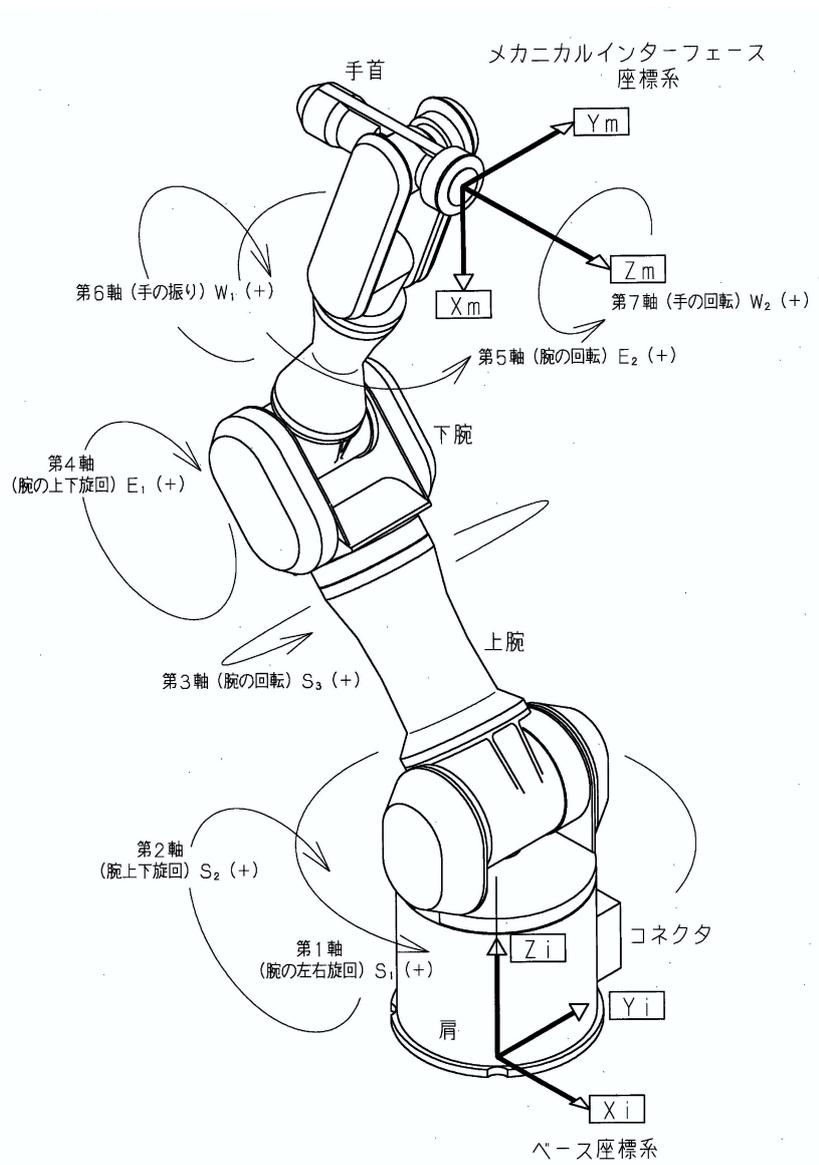


図 44 マニピュレータ PA10 の回転軸

C. 6軸マニピュレータの逆運動学

マニピュレータのリンク構造を図 45 に示す . 7 軸マニピュレータ PA10 の 3 軸目 Joint3 を固定し , PUMA 型の 6 軸マニピュレータとして利用した .

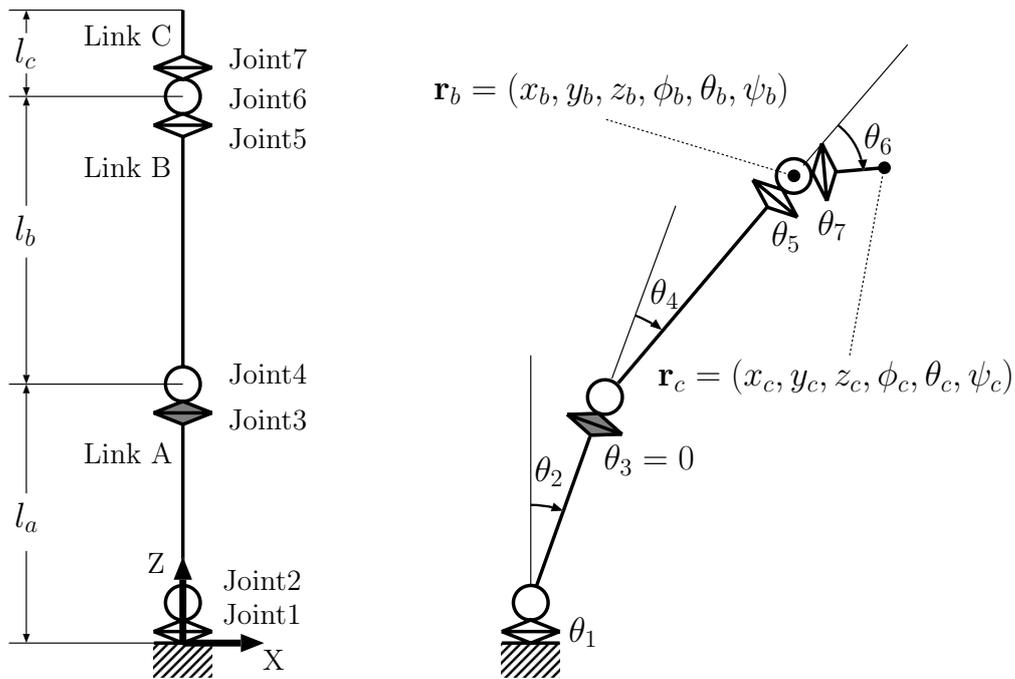


図 45 第 3 関節を固定したマニピュレータ PA10 の関節構造

このリンク構造では Joint5 から Joint7 の 3 軸の原点が全て Joint6 上で交わるため , 手先の位置 , 姿勢 (ロールピッチヨー表現) r_c が得られたとすると , 3 軸の原点 r_b の位置 x_b, y_b, z_b は一意に決定される .

$$x_b = x_c - l_c \cos \phi_c \sin \theta_c \quad (9)$$

$$y_b = y_c - l_c \sin \phi_c \sin \theta_c \quad (10)$$

$$z_b = z_c - l_c \cos \theta_c \quad (11)$$

この原点位置 x_b, y_b, z_b は, Joint1, Joint2, Joint4 の関節角 $\theta_1, \theta_2, \theta_4$ で表現できるので, $\theta_1, \theta_2, \theta_4$ は x_b, y_b, z_b を用いて以下の様に表せる.

$$\theta_1 = \text{atan2}(y_b, x_b) \quad (12)$$

$$\theta_2 = \cos^{-1} \frac{z_b}{\sqrt{x_b^2 + y_b^2 + z_b^2}} - \cos^{-1} \frac{l_a^2 - l_b^2 + x_b^2 + y_b^2 + z_b^2}{2l_a \sqrt{x_b^2 + y_b^2 + z_b^2}} \quad (13)$$

$$\theta_4 = \cos^{-1} \frac{x_b^2 + y_b^2 + z_b^2 - l_a^2 - l_b^2}{2l_a l_b} \quad (14)$$

ただし,

$$\sqrt{x_b^2 + y_b^2 + z_b^2} < l_a + l_b \quad (15)$$

根元3軸の関節角 $\theta_1, \theta_2, \theta_4$ より, 順運動学を用いて原点 r_b の姿勢 $(\phi_b, \theta_b, \psi_b)$ が決定される. 最後に, 原点 r_b から見た原点 r_c の相対的な姿勢をオイラー表現で表すことで, 手先の3軸の関節角 $\theta_5, \theta_6, \theta_7$ を決定できる.

D. 回転行列から回転要素の導出

D.1 ロールピッチヨー角への変換式

ロールピッチヨー表現での回転行列を以下のようにおく.

$$\begin{aligned} \mathbf{R} &= \mathbf{R}(Z, \phi) \mathbf{R}(Y, \theta) \mathbf{R}(X, \psi) \\ &= \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{bmatrix} \\ &= \begin{bmatrix} \cos \phi \cos \theta & \cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi & \cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi \\ \sin \phi \cos \theta & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi & \sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi \\ -\sin \theta & \cos \theta \sin \psi & \cos \theta \cos \psi \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (16)$$

式 (16) において，各要素を取り出すと以下の方程式が得られる．

$$\cos \phi \cos \theta = R_{11} \quad (17)$$

$$\cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi = R_{12} \quad (18)$$

$$\cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi = R_{13} \quad (19)$$

$$\sin \phi \cos \theta = R_{21} \quad (20)$$

$$\sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi = R_{22} \quad (21)$$

$$\sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi = R_{23} \quad (22)$$

$$-\sin \theta = R_{31} \quad (23)$$

$$\cos \theta \sin \psi = R_{32} \quad (24)$$

$$\cos \theta \cos \psi = R_{33} \quad (25)$$

式 (24)，式 (25) より

$$\begin{aligned} \cos^2 \theta (\sin^2 \psi + \cos^2 \psi) &= R_{32}^2 + R_{33}^2 \\ \cos \theta &= \pm \sqrt{R_{32}^2 + R_{33}^2} \end{aligned} \quad (26)$$

もし $\cos \theta \neq 0$ つまり $\theta \neq \pm 90^\circ$ ならば，式 (23)，式 (26) より

$$\begin{aligned} \frac{-\sin \theta}{\cos \theta} &= \frac{R_{31}}{\pm \sqrt{R_{32}^2 + R_{33}^2}} \\ \tan \theta &= \frac{-R_{31}}{\pm \sqrt{R_{32}^2 + R_{33}^2}} \\ \theta &= \text{atan2}(\mp R_{31}, \pm \sqrt{R_{32}^2 + R_{33}^2}) \end{aligned} \quad (27)$$

式 (24)，式 (25) より

$$\begin{aligned} \frac{\sin \psi}{\cos \psi} &= \frac{R_{32}}{R_{33}} \\ \psi &= \text{atan2}(\pm R_{32}, \pm R_{33}) \end{aligned} \quad (28)$$

式 (17), 式 (20) より

$$\frac{\sin \phi}{\cos \phi} = \frac{R_{21}}{R_{11}}$$

$$\phi = \text{atan2}(\pm R_{21}, \pm R_{11}) \quad (29)$$

したがって, $R_{32}^2 + R_{33}^2 \neq 0$ ならロールピッチヨー角は式 (27), 式 (28), 式 (29) で与えられる. つまり回転行列 R が与えられたとき式 (16) を満足するロールピッチヨー角は2通り存在する. なお, θ を $-90^\circ < \theta < 90^\circ$ の範囲に限定すると,

$$\phi = \text{atan2}(R_{21}, R_{11}) \quad (30)$$

$$\theta = \text{atan2}(-R_{31}, \sqrt{R_{32}^2 + R_{33}^2}) \quad (31)$$

$$\psi = \text{atan2}(R_{32}, R_{33}) \quad (32)$$

となり, ロールピッチヨー角は一意に定まる. また $\cos \theta = 0$ すなわち $R_{32}^2 + R_{33}^2 = 0$ ならば,

$$\phi = \text{任意} \quad (33)$$

$$\theta = 90^\circ \text{または} -90^\circ \quad (34)$$

$$\psi = -\text{atan2}(R_{23}, R_{22}) \quad (35)$$

D.2 オイラー角への変換式

回転行列からオイラー角への変換式の導出は参考文献 [50] に詳しく記述されており, こちらを参照されたい.

オイラー表現での回転行列を以下のようにおくと,

$$\begin{aligned} \mathbf{R} &= \mathbf{R}(Z, \phi) \mathbf{R}(Y, \theta) \mathbf{R}(Z, \psi) \\ &= \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} \cos \phi \cos \theta \cos \psi - \sin \phi \sin \psi & -\cos \phi \cos \theta \sin \psi - \sin \phi \cos \psi & \cos \phi \sin \theta \\ \sin \phi \cos \theta \cos \psi + \cos \phi \sin \psi & -\sin \phi \cos \theta \sin \psi + \cos \phi \cos \psi & \sin \phi \sin \theta \\ & -\sin \theta \cos \psi & \sin \theta \sin \psi & \cos \theta \end{bmatrix} \\
&= \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \tag{36}
\end{aligned}$$

オイラー角は θ を $0^\circ < \theta < 180^\circ$ の範囲に限定すると，以下のように表される．

$$\phi = \text{atan2}(R_{23}, R_{13}) \tag{37}$$

$$\theta = \text{atan2}(\sqrt{R_{13}^2 + R_{23}^2}, R_{33}) \tag{38}$$

$$\psi = \text{atan2}(R_{32}, -R_{31}) \tag{39}$$

また， $R_{13}^2 + R_{23}^2 = 0$ ならば，

$$\phi = \text{任意} \tag{40}$$

$$\theta = 90^\circ - R_{33} \times 90^\circ \tag{41}$$

$$\psi = \text{atan2}(R_{21}, R_{22}) - R_{33}\phi \tag{42}$$