

筑波大学大学院博士課程

システム情報工学研究科修士論文

HMD 及び大画面環境に適したハンドジェスチャ
インタラクション

DAVAASUREN ENKHBAT

修士（工学）

（コンピュータサイエンス専攻）

指導教員 田中 二郎

2014年3月

概要

視覚的に情報を得るためのデバイスとしてのディスプレイが日々進化しており、小型で持ち運ぶことができるディスプレイから巨大な大画面まで様々な種類のディスプレイが存在している。特に、近年 HMD（頭部装着型表示デバイスであり、本文で詳しく説明する）がより小型で使いやすいウェアラブルディスプレイとして世間の注目を浴びており、大手 IT 企業たちも大きな可能性を持つ新しい分野として、HMD の研究開発に参入している。本研究では、この状況が HMD の進化を更に加速させ、近い将来人々が日常生活の中で利用する時代が到来すると想定している。特に、HMD が個人用モバイルデバイスとして使われ、人々がそこに表示される個人情報と外部のディスプレイ上の共有情報を使い分けながら利用することが考えられる。

本研究では、このように HMD が個人用モバイルデバイス、外部の大画面が情報を共有するためのデバイスとして使われる環境でこれらのデバイスとの直感的なインタラクションとデバイス間での情報交換などが重要なテーマであると考えた。そして、HMD 及び大画面のインタラクション手法として従来から研究されているハンドジェスチャインタラクション手法に注目し、これらのデバイスが混在した未来環境におけるハンドジェスチャインタラクションの問題点を解決したインタラクション手法の提案、及びプロトタイプシステムの開発を行った。

プロトタイプシステムでは HMD に取り付けられた RGB 及び深度カメラを用いてユーザのハンドジェスチャを認識し、HMD 及び大画面でインタラクションを行うことができる。また、ユーザが手の左右や手のひらの向きを変える簡単なハンドジェスチャでインタラクション対象を切り替えることができる。

目次

第1章	序論	1
1.1	プライベート画面インタフェース	1
1.2	パブリック画面インタフェース	3
1.3	未来環境の予想	3
1.4	HMD及び大画面環境におけるインタラクション手法	3
1.4.1	ハンドジェスチャインタラクション	4
1.5	未来環境における従来ハンドジェスチャの問題点	5
1.6	本研究の目的	6
1.7	本研究のアプローチ	7
1.8	本論文の構成	7
第2章	HMD及び大画面環境に適したハンドジェスチャインタラクションシステム	8
2.1	想定する未来環境	9
2.2	提案インタラクション	9
2.3	利用シナリオ	12
2.3.1	会議室での利用例	12
2.3.2	写真共有場の例	13
第3章	プロトタイプシステム	14
3.1	システム概要	14
3.2	インタラクション	14
3.2.1	プライベート画面上の画像に対するインタラクション	15
	画像の表示・非表示	15
	画像の切り替え	15
	画像の拡大縮小	16
	画像の回転	16
3.2.2	パブリック画面上の画像に対するインタラクション	17
	画像の配置変更	17
	画像の回転, 拡大縮小	17
3.2.3	プライベート及びパブリック画面間での情報共有	18
	画像の提示	18
	画像の取得	18

第4章	実装	19
4.1	開発環境及びシステム構成	19
4.2	通信方式	20
4.3	大画面の識別及び位置関係	21
4.3.1	大画面の識別	21
4.3.2	大画面の位置及び向き	21
4.4	アプリケーションの実装	22
4.5	ハンドジェスチャの認識	23
4.5.1	両手の領域の検出	23
4.5.2	各領域の左手か右手かの判別	24
4.5.3	手の軸の検出	25
4.5.4	指先の検出	25
4.5.5	手のひらの向きの判別	26
	親指の検出	26
	手の軸と親指の位置関係	26
4.5.6	ジェスチャ判別	27
4.6	指で突き抜く動作の判定及びオクルージョンの再現	27
4.6.1	突き抜く動作の判定	27
4.6.2	オクルージョンの再現	28
第5章	関連研究	29
5.1	HMDインタラクションに関する研究	29
5.2	大画面インタラクションに関する研究	30
第6章	課題と今後の発展	31
6.1	課題	31
6.2	今後の発展	31
第7章	結論	32
	謝辞	33
	参考文献	34

目次

1.1	透過型 HMD	2
1.2	音声によるインタラクションの問題	4
1.3	ハンドジェスチャインタラクション (先行研究 [8] からの写真)	5
2.1	提案システムのイメージ	8
2.2	インタラクション対象の切り替え	10
2.3	画面インタフェースでのインタラクション	11
2.4	会議室での利用例	12
2.5	写真共有場の例	13
3.1	プロトタイプシステムの装着様子	14
3.2	画像の表示・非表示	15
3.3	画像の切り替え	15
3.4	画像の拡大縮小	16
3.5	画像の回転	16
3.6	パブリック画面の GUI	17
3.7	画像の位置変更及び回転, 拡大縮小	17
3.8	画像の共有と取得	18
4.1	システム構成	19
4.2	通信方式	20
4.3	大画面の特定	21
4.4	カメラ座標系とマーカ座標系	21
4.5	大画面上でのハンドジェスチャのフィードバック	22
4.6	DepthMap 及び ConfidenceMap	23
4.7	手のマスク画像及び手の深度マップ	24
4.8	手の左右の判別	24
4.9	手の先端及び手の軸の検出	25
4.10	指先検出	26
4.11	親指の検出と手の左右の判別	26
4.12	片手のハンドジェスチャ	27
4.13	オクルージョンの再現	28

表目次

2.1 左手のひらの向きによるインタラクション対象の切り替え	9
2.2 画面インタフェースでのインタラクション	11

第1章 序論

現在人々は、情報を重要な資源とする超情報化社会に突入したといわれている。IT技術の飛躍的進歩やインターネットという全世界を結ぶコンピュータネットワークによって、誰もが自由かつ簡単に情報を作り、発信し、誰かと共有し、あるいは情報を取得することが可能になった。そのため、情報が急激に増大し、人々が莫大な情報に触れながら生活しなければならなくなった。そしてその大量な情報の中から自分に有益な情報をいかに効率よく見つけ出すかが人々の重要な課題になったといえる。

人間は五感を通じて外部から情報を取得するが、その8割以上を視覚によって得られるといわれており、人間にとって視覚による情報伝達が最も効率的な手段であると考えられる。そのため、以前からディスプレイは情報を視覚的に提示するデバイスとして重要視され、改良されてきた。現在人々が、携帯端末やスマートフォンの画面やパソコンのディスプレイ、テレビ、会議室の大画面、プロジェクタなど様々なディスプレイを日常的にあらゆる場面や用途で一般に使っており、持ち運び安い小型ディスプレイから高解像度で大きなディスプレイまで様々な種類のディスプレイが存在している。そして、ディスプレイに関する技術も日々進化し続けており、近年その中で特にHMDと大画面が世間の注目を浴びている。本研究では、HMDと大画面の進化によって人々の情報を扱う環境が大きく変わると想定している。

HMDとは

Head Mounted Displayの略語であり、頭部に装着するディスプレイのことである(図1.1)。ディスプレイ部分が常にユーザの目の近くにあるためユーザの姿勢や動作などに影響を受けずに映像を表示できるのが特徴である。そしてHMDは、ユーザの視覚情報を完全に遮断し、コンピュータによる映像情報のみ提示する非透過型と、カメラあるいはハーフミラーを用いて視覚情報を提示しながら、コンピュータによる付加情報を提示する透過型の2種類に大きく分けられている。特に、ビデオ透過型と呼ばれる、非透過型HMDにカメラを取り付けたものが研究材料として注目されている。

1.1 プライベート画面インタフェース

HMDの用途としては、非透過型HMDは以前からVR(仮想現実)に関する研究分野で、透過型HMDは、AR(拡張現実)という現実の情報をコンピュータによる付加情報によって拡張する研究分野で研究材料として盛んに使われ、開発が進められてきた。その一方で、一般には業務用や軍用など特定の用途にのみ使われており、社会にはほとんど普及していなかった。その理由として、小型に作っているためユーザに提示できる映像の解像度が低く、ディス

プレイサイズも小さい一方で、実際に装着して行動するには大きくて、重いというハードウェア的問題や、装着すると外見が目立ったり、目がディスプレイで覆われてしまうため相手に目が見なく、表情が相手に伝わりにくくなったりといった人間同士のコミュニケーションや社会に簡単に受け入れられないという問題が挙げられる。

しかし、近年その状況が一変し、HMD が世間に大きく取り上げられるようになり、次世代のウェアブルデバイスとして注目を浴びている。そのきっかけの一つは大手 IT 企業である Google 社が Google Glass¹ (図 1.1 - A) という独自に開発した透過型 HMD を世間に発表したことである。Google 社の携帯端末 OS である Android² を搭載しており、インターネット接続や音声認識、カメラ機能が完備されている。これに続き、他の大手 IT 関連企業たちが独自に開発した透過型 HMD の試作や商品 (Vuzix 社の M-100³ (図 1.1 - B), Epson 社の Moverio BT-100⁴ (図 1.1 - C), Scalar 社の Teleglass T3⁵ (図 1.1 - D), Telepathy 社の Telepathy One⁶ (図 1.1 - E)) が次々と発表されている。これらの HMD は、従来の HMD より軽量化、小型化されていて、装着していても比較的目立たなくなっているためより社会に受け入れやすいデバイスになっている。現在はまだ表示領域が小さいが、将来的に目の視界を妨げない大きさになることが予想され、現在画面の小ささが問題視されるスマートフォンが HMD によって置き換えられる時代が来ると考えられる。非透過型 HMD についても、Sony 社が開発した高解像度 HMD である HMZ-T2⁷ や広い視界を実現した OculusVR 社の OculusRift⁸ などが新しく発表され、映像鑑賞やゲーム用のデバイスとして注目浴びている。



図 1.1: 透過型 HMD

本研究では HMD を取り巻くこのような流れが今後も続き、HMD が持つ従来の問題が解決されることによって、世間一般に普及すると予想している。人々が HMD を用いて映像や写

¹<http://www.google.com/glass/start/>

²<http://www.android.com/>

³http://www.vuzix.com/consumer/products_m100/

⁴<http://www.epson.jp/products/moverio/>

⁵<http://www.scalar.co.jp/products/t3.php>

⁶<http://tele-path.org/>

⁷<http://www.sony.jp/hmd/products/HMZ-T2/>

⁸<http://www.oculusvr.com/>

真、音楽を楽しんだり、電子本や新聞を読んだり、インターネットで家族や友達と電話したり、ゲームをしたりなど生活の中で常時装着して使うプライベート画面インタフェースとして利用することが予想される。しかも、場所や姿勢など空間的制約にとらわれず好きな場所で利用できるため人々の生活の利便性に繋がると考えられる。

1.2 パブリック画面インタフェース

HMDのようにディスプレイの小型化、軽量化が進む一方でより高解像度でより大きなディスプレイも近年増えており、家庭におけるテレビとしてももちろん、学校や職場における会議用の大画面としてや、表示する内容を自由に変更できることから博物館、デパートメントストア、駅など共用スペースにおける不特定多数のユーザ向けディスプレイとしてよく利用されている。また、最近では自動販売機にも使われている例もあり、単なる表示デバイスとしてではなく、インタラクション可能なデバイスとしても今後使われていくことが予想される。特に、人々が情報を載せたり、あるいはそこから取得したりするなど誰もが自由にアクセスできるインタラクティブなディスプレイとしての利用が増えると考えられる。

本研究では、大画面が情報を共有するためのパブリック画面インタフェースとしてよく使われ、人々がそれを通して繋がり、共同作業を行うことで人々のコミュニケーションがより活発で円滑に行われるようになると予想している。

また、パブリック画面インタフェースには物理的な大画面に限らず加茂ら [10] や金子ら [11] よって提案されたような HMD を装着した複数のユーザが共有して使う仮想ディスプレイインタフェースも含まれるが、本研究では物理的な大画面をパブリック画面インタフェースとして扱う。

1.3 未来環境の予想

上に説明したようにプライベート及びパブリック画面インタフェースが混在する環境が社会に普及し、人々が日常生活の中でそれらとインタラクションを行いながら、生活する環境が近い将来実現されることが予想される。

1.4 HMD 及び大画面環境におけるインタラクション手法

本研究では、HMD によるプライベート画面及び、大画面によるパブリック画面が共存する環境においてインタラクションを行う際、マウスやキーボード、タッチ入力、音声入力など従来のインタラクション手法では操作することが困難になると考えている。そこで本研究では、ジェスチャインタラクションがこれらの画面インタフェースとのインタラクションに最も有効だと考え、注目した。

例えば、HMD においてマウスやキーボードを使うことが HMD の携帯性を損なわれてしまい、持ち運びに向かない。また、MOVERIO BT-100 などで使われているような手に持って操

作するタッチパネル付きデバイスを使う方法もあるが、表示と操作が別々になってしまうため操作性が失われてしまう問題があった。また、Google Glass ような音声認識によるインタラクション方法はテキスト入力や抽象的コマンド操作の際に有効だが、使う場所のマナーやプライバシーの問題が挙げられる。例えば、静かな公共の場では音声操作が不適切である（図 1.2）。



図 1.2: 音声によるインタラクションの問題

大画面に関しても、マウスやキーボード、タッチ入力によるインタラクションは画面の大きさや大画面との距離によって使えない可能性がある。音声認識によるインタラクションも HMD の場合と同様に困難な場合が考えられる。

一方でジェスチャインタラクションは、より直観的なインタラクション手法として近年盛んに研究開発が行われ、次世代のインタラクション手法として注目浴びている手法である。また、マウスやキーボードなどユーザの携帯性を損なわせるデバイスを用いないインタラクション手法であるため本研究ではジェスチャインタラクションに注目した。

1.4.1 ハンドジェスチャインタラクション

ハンドジェスチャインタラクションとは、ユーザが体の身体動作でコンピュータとインタラクションできるジェスチャインタラクションの内、手の姿勢や動作などのジェスチャを用いてインタラクションを行う手法のことであり、ユーザの手の自然な動作をコンピュータの操作に割り与えることで直感的なインタラクションが可能なる（図 1.3）。

ハンドジェスチャインタラクションはマウスやキーボード、タッチ入力、音声入力に比べて、HMD の携帯性を損なわず、場所に関係なく操作することが可能であり、また、大画面の大きさや距離によって操作性が失われないというメリットがある。そのため、HMD 及び大画面とのインタラクション手法としてハンドジェスチャインタラクションが最も有効だと考えられる。

ハンドジェスチャインタラクションを実現するための方法としてユーザの手にセンサを取り付けることで手の姿勢や動きを認識するセンサベース手法と、ユーザあるいは空間に取り



図 1.3: ハンドジェスチャインタラクション (先行研究 [8] からの写真)

付けたカメラによってユーザの手の姿勢や動きを認識するビジョンベース手法がある。センサベース手法は手の状態を表す信号をセンサによって直接取得できるので雑音が多く、高い認識精度を得られる。その反面、インタラクションの際にユーザが常に手に装置を付けて行動しないとけなく、携帯性の面で問題があった。ビジョンベース手法は手に何も付ける必要がなく、携帯性に優れているが、カメラによるキャプチャ画像を解析して手を認識する際にカメラの視界内ではしか認識できないことや、周辺環境の照明に影響されやすいなど認識精度に問題があった。またカメラの設置場所によって携帯性も損なわせる場合も考えられる。しかし近年、Kinect⁹をはじめとする被写体との距離がわかる深度カメラの出現により、身体の動作をより高い認識精度で得られるようになった。また LeapMotion¹⁰、DepthSense¹¹ など小型で持ち運びやすい深度カメラも登場しており、携帯性を損なわずにハンドジェスチャインタラクションを実現可能になった。

1.5 未来環境における従来ハンドジェスチャの問題点

ハンドジェスチャインタラクションは HMD、あるいは大画面とのインタラクション手法として有効だが、HMD と大画面が混在する未来環境において現在使われているハンドジェスチャインタラクションでは問題があると考えられる。

従来のハンドジェスチャインタラクションは、HMD 用インタラクションがユーザ側で、大画面用インタラクションが大画面側で認識及び管理される HMD と大画面のそれぞれに特化したシステムの場合が多い。このように HMD と大画面で別々に認識、管理される手法では未来環境において次のような問題が考えられる。

⁹<http://www.microsoft.com/en-us/kinectforwindows/>

¹⁰<https://www.leapmotion.com/>

¹¹<http://www.softkinetic.com/products/depthsensecameras.aspx>

システムパフォーマンスの問題

ユーザの安定した正確な識別が画面インタフェース間での情報共有の際に欠かせない重要な問題である。大画面に取り付けたカメラを用いてハンドジェスチャを認識する従来手法ではユーザの位置が簡単に分かるが、ユーザの識別にはユーザに取り付けたマーカやユーザの顔の特徴などを用いる手法がよく使われている。しかし、複数人で複数の大画面とインタラクションを行う未来環境において、大画面ごとに全てのユーザのハンドジェスチャ及びユーザの識別を行うとなると、各大画面における処理が増え、システム全体のパフォーマンスに影響が出る問題が考えられる。

ユーザやそのハンドジェスチャの認識の問題

複数人でインタラクションを行う際に、ユーザが大画面から離れた位置でインタラクションを行う場面や、ユーザ同士が近い位置でインタラクションを行う場面が考えられる。その際に大画面に取り付けたカメラで認識する従来のインタラクション手法では、遠くにいるユーザを識別できなかったり、ユーザの細かいハンドジェスチャを認識できなかったりする問題が考えられる。また、近くにいるユーザ同士が大画面に対して前後に重なってしまい、後ろのユーザやそのハンドジェスチャが正しく認識されないという問題が考えられる。つまり、大画面のカメラの視界内に全身が鮮明に映っているユーザしかインタラクションを行うことができないという問題が考えられる。将来カメラの性能が向上し、視界が広がったとしてもこの問題の解決にはならない。

インタラクション対象の切り替えの問題

複数の HMD 及び大画面が混在する未来環境では、それらのデバイス間での頻繁な情報の通信を行う場面が考えられる。その際に、ユーザが接続させたいデバイスを切り替えるたびに手動でデバイスを指定し、接続操作を行うというインタラクションと関係ない動作に時間をかけてしまい、スムーズなインタラクションの妨げになるという問題が考えられる。

誤認識、誤操作の問題

HMD と大画面が混在している環境で従来の手法のように HMD と大画面がそれぞれのカメラでユーザのハンドジェスチャを認識すると、ユーザが HMD を操作しようとして、周囲の大画面で意図しない誤操作を行ってしまう可能性が考えられる。つまり、ハンドジェスチャを HMD と大画面で 2 重に認識されてしまうという問題が考えられる。また、画面インタラクションを目的としなかった手を振る、指さすなどの日常的な動作が誤操作を引き起こす可能性も考えられる。

1.6 本研究の目的

本研究の目的は、人々が日常生活の中で HMD を持ち運ぶプライベート画面インタフェースとして使い、あらゆる場所に設置された大画面を情報共有のパブリック画面インタフェース

スとして利用する未来環境に適した、前述した従来のハンドジェスチャインタラクションの問題を解決できるインタラクションシステムを実現することである。

具体的には以下の4つの要件を満たすシステムを実現する。

- システム全体のパフォーマンスがユーザの人数に影響を受けないシステム。
- 携帯性を損なわず、ユーザやそのハンドジェスチャを正確に認識できるシステム。
- 画面インタフェース間で動的に切り替えてインタラクションできるシステム。
- インタラクションの際に誤認識、誤操作が起こりにくいシステム。

1.7 本研究のアプローチ

これらの要件を満たすためのアプローチとして、大画面に対するハンドジェスチャを HMD 側で認識し、ネットワークを通して大画面に伝えることでインタラクションを行うようにする。このようにそれぞれのデバイスのみでユーザのハンドジェスチャを認識するため、システム全体の処理がデバイスことに分散され、ユーザの人数に影響を受けないシステムを作ることができる。ユーザ同士もお互いの妨げにならずにハンドジェスチャインタラクションすることができる。

ユーザの識別もデバイスを識別することで容易になり、ユーザをその携帯性を損なわずに正確に識別することができる。また、このように各ユーザのハンドジェスチャがユーザの端末のみで認識されるため重複認識による誤操作の問題も解決されることが考えられる。

また、インタラクションの際に手の左右や手のひらの向きを用いることで HMD に対するインタラクション、及び大画面に対するインタラクションを開始や停止させたり、画面間で切り替えられたりできるようにする。これによってシステムにおける誤動作や誤操作を防ぐことができると考えられる。

そして、大画面側にマーカを取り付け、それを HMD に取り付けた RGB カメラで読み取ることでシステムが動的に大画面を特定し、接続できるシステムを実現する。

1.8 本論文の構成

本論文は、まず始めに第2章で本研究で提案する HMD 及び大画面環境に適したハンドジェスチャインタラクションシステムについて説明する。そして、第3章では今回実装を行ったプロトタイプシステムの概要やそのインタラクションについて説明する。第4章では、プロトタイプシステムの実装方法について述べる。第5章で本研究の関連研究について述べ、研究の位置を行う。第6章では今後の課題と将来の発展について述べ、最後の第7章にて結論を述べる。

第2章 HMD 及び大画面環境に適したハンド ジェスチャインタラクションシステム

本システムはカメラ付き HMD を装着したユーザが HMD 内に映るプライベート画面，及び周辺の空間に固定されたパブリック画面と両手のハンドジェスチャによってインタラクションを行うことができるシステムであり（図 2.1），以下のようなメリットがある．

- ユーザのハンドジェスチャは全てユーザの HMD で認識されるためユーザは場所に拘束されることなくインタラクションを行うことができる．
- ユーザが複数人いても，それぞれの HMD でインタラクションを行うため，システム全体にかかる負担も少なく，ユーザ同士の干渉問題が起きにくい．
- 直感的なハンドジェスチャでインタラクション対象をプライベート画面あるいはパブリック画面に簡単に切り替えることができる．

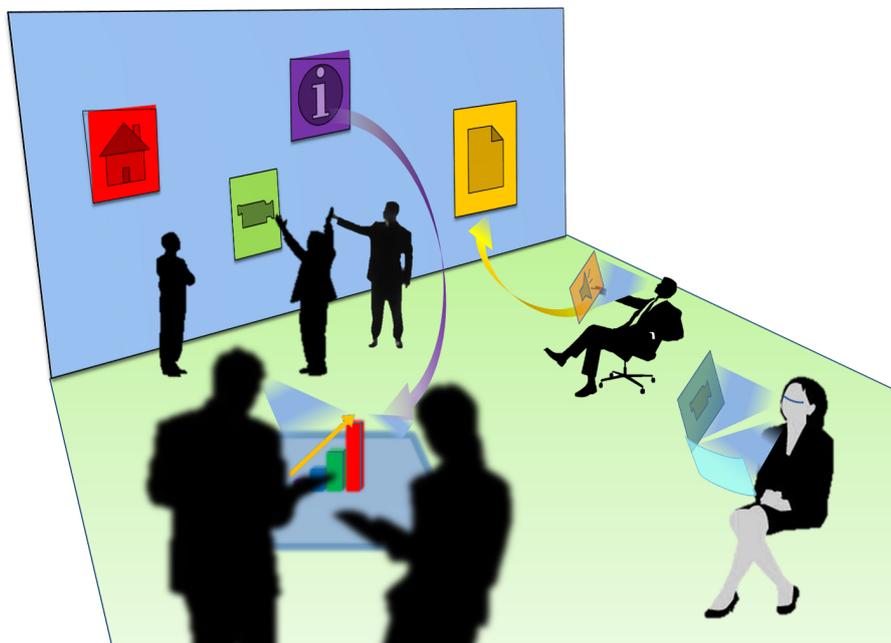


図 2.1: 提案システムのイメージ

2.1 想定する未来環境

本研究では以下のような未来環境を想定する。

- HMD の進歩が進み個人ウェアブルデバイスとして社会に広く普及している。
- あらゆる共用スペースにインタラクティブな大画面が設置され、人々が情報共有の場として使う大画面環境が整っている。
- 人々がハンドジェスチャを使い、日常生活の中でこれらのデバイスと頻繁にインタラクションを行っている。

2.2 提案インタラクション

本システムでは、インタラクションの方針として、コンテンツの細かい操作を利き手のジェスチャで行い、その他の画面の切り替えなど画面コントロールを非利き手のジェスチャで行う。また、ユーザが画面インタフェースをわかりやすく使い分けられるように、プライベート画面では手のひらを自分に向けたジェスチャ、パブリック画面では手のひらを外に向けたジェスチャを用いている。そして、一般的に右利きのユーザを想定し、以下のようなハンドジェスチャインタラクションを提案する。

まず、プライベート画面とパブリック画面のコントロールは以下のようなハンドジェスチャで行う（表 2.1）。

表 2.1: 左手のひらの向きによるインタラクション対象の切り替え

ジェスチャ	インタラクション
手のひらを自分に向けた5本指のジェスチャ（図 2.2 の上）	プライベート画面を表示，あるいは非表示させる。プライベート画面を表示中はパブリック画面がインタラクション不可になる。
手のひらを外に向けた5本指のジェスチャ（図 2.2 の下）	プライベート画面が非表示されているときに，パブリック画面とのインタラクションを開始，あるいは停止させる。

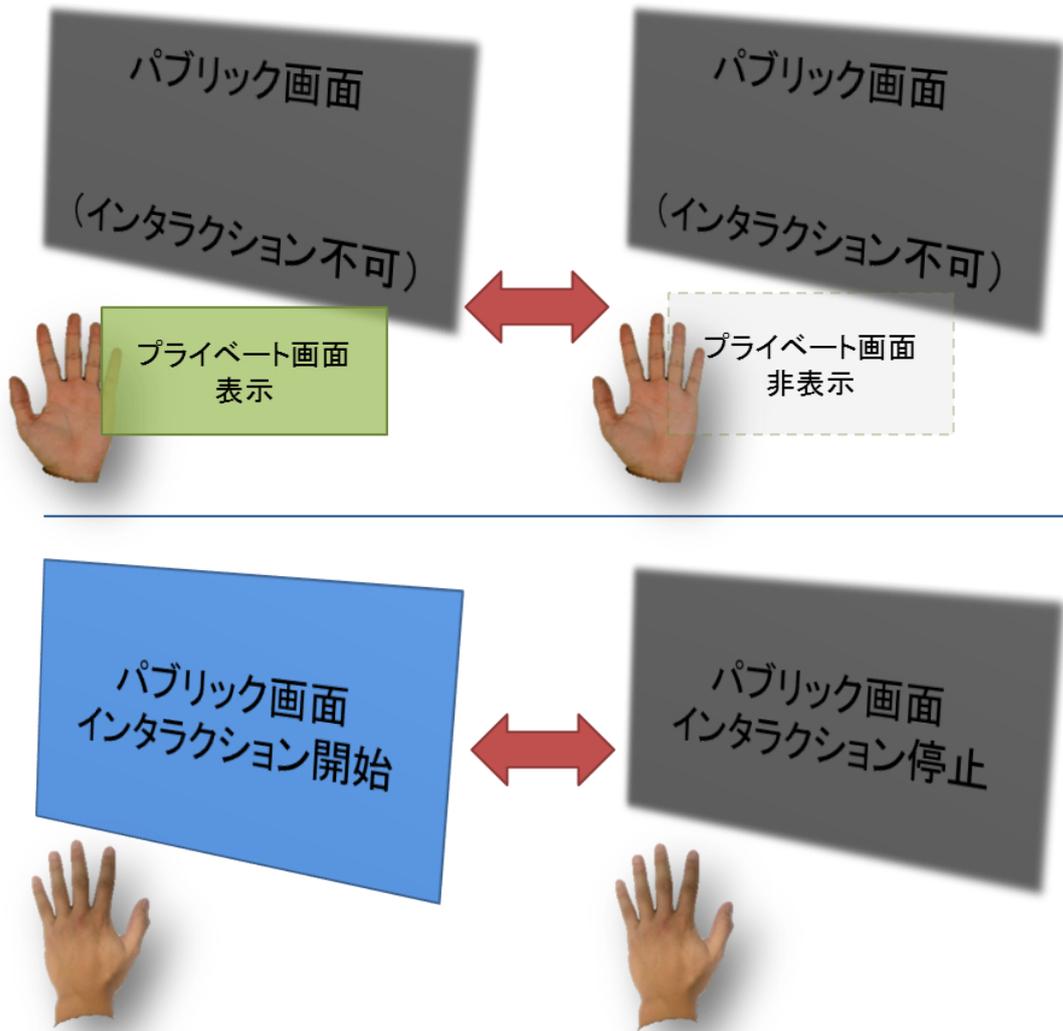


図 2.2: インタラクション対象の切り替え

そして、プライベート画面，あるいはパブリック画面上での情報に対して以下のようなジェスチャインタラクションを行うことができる（表 2.2）．これらはユーザが覚えやすく，より直感的なインタラクションができるようにタッチ入力スマートフォンやタブレット端末などで現在よく使われている直接操作のジェスチャを参考にしている．

表 2.2: 画面インタフェースでのインタラクション

ジェスチャ	インタラクション
右手の1本指でのタップジェスチャ（図 2.3-A）	タップして一次選択する
右手の1本指でのフリックジェスチャ（図 2.3-B）	次または前に切り替える
右手の1本指でのロングタップジェスチャ（図 2.3-C）	プライベート画面とパブリック画面間で移動させる
左手の4本指のジェスチャと右手の1本指のなぞるジェスチャ（図 2.3-D）	画面上で手書きする
右手の2本指でのフリックジェスチャ（図 2.3-E）	画面内でドラッグする
右手の2本指でのピンチジェスチャ（図 2.3-F）	表示サイズを変更する
右手の2本指での回転ジェスチャ（図 2.3-G）	傾きを変更する
両手の2本指でのフォーカスジェスチャ（図 2.3-H）	HMD のカメラで写真を撮る

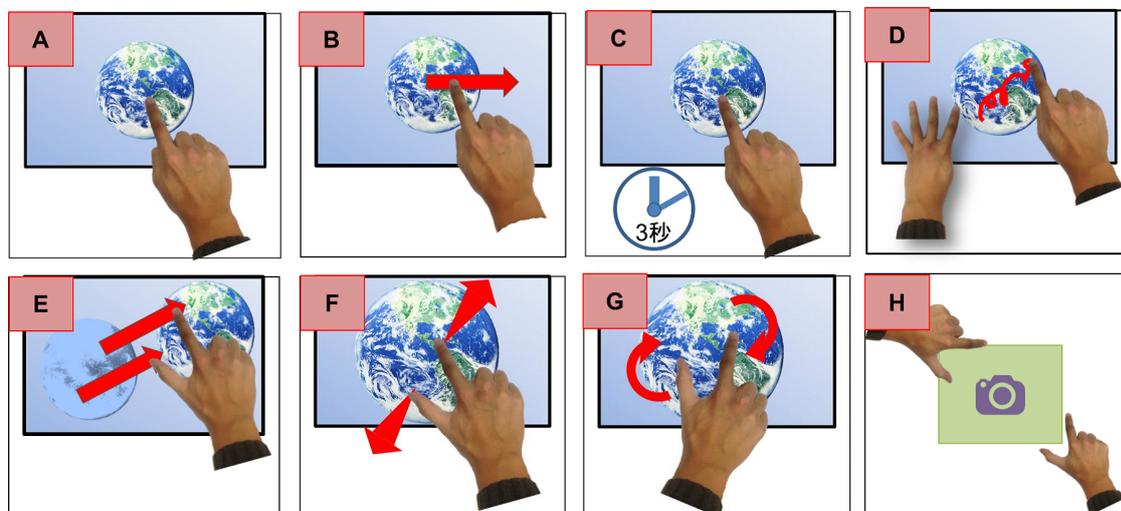


図 2.3: 画面インタフェースでのインタラクション

2.3 利用シナリオ

2.3.1 会議室での利用例



図 2.4: 会議室での利用例

Aさんが自分の発表資料を自分のプライベート画面上で確認してから、ハンドジェスチャで会議室のパブリック画面上に素早く共有した。発表を聞きに来た研究室の学生たちもハンドジェスチャでAさんの資料を素早く自分のプライベート画面で表示させた。Aさんが発表中に自分のプライベート画面で発表のメモを確認しながら重要なポイントを逃さずにスムーズに発表を行った。また、発表中にデバイスの操作を気にせず、スライドを送る自然なハンドジェスチャのみで発表を進めた。それと同時に発表を聞いていた学生たちがそれぞれのプライベート画面で資料を確認しながら、ハンドジェスチャで気になるところにマークや文字を書き加えたり、スライドの図を拡大して詳しく見てみたりした。そして、質疑応答の際にその情報をパブリック画面に共有しながら話し合うことができた。その結果、情報交換がより効率的になり、より活発な議論が行われた(図 2.4)。

2.3.2 写真共有場の例



図 2.5: 写真共有場の例

Eさんが職場のパブリック画面の前にやって来た。そこには同僚たちも集まっており、新年会のパーティで撮った写真や動画がみんなで共有しようとそれぞれの端末から大画面上に情報を載せ、写真や動画をの話題で盛り上がっていた。Eさんが、気になる写真が載っているのを見てジェスチャで自分の端末に素早くコピーできた。そして、自分で撮った写真や動画をプライベート画面上で確認しながら共有したいと思った写真や動画をハンドジェスチャで素早くパブリック画面上に共有できた（図 2.5）。

これらの利用シナリオでは、大画面の前に集まった複数のユーザのハンドジェスチャはそれぞれの HMD で認識、管理されているためユーザの認識や識別が簡単に行われ、ユーザ同士がお互いに干渉し合わずにインタラクションを行うことができている。

そして、ユーザが自分にだけ見えるプライベート画面と他のユーザにも見えるパブリック画面を使い分けながら、ハンドジェスチャで直感的にインタラクションを行い、画面間で情報を移動させることもできている。

第3章 プロトタイプシステム

3.1 システム概要

本研究では提案システムのプロトタイプシステムとしてハンドジェスチャのみでインタラクションを行うことができる簡単な画像共有システムを開発した(図3.1)。まず、プライベート画面インタフェースのアプリケーションとしてユーザが目の前で画像情報を閲覧させ、拡大縮小、回転、スライドなどのインタラクションを行うことができるアプリケーションを実装した。また、画像を切り替えたり、選択してパブリック画面インタフェースに共有したりすることができる。パブリック画面インタフェースのアプリケーションも画像情報データを提示、移動、回転、拡大縮小などプライベート画面と同様なインタラクションが可能なアプリケーションを実装した。



図 3.1: プロトタイプシステムの装着様子

3.2 インタラクション

本システムではプライベート及びパブリック画面に対して以下のようなインタラクションが可能であり、それらについて詳しく説明する。

3.2.1 プライベート画面上の画像に対するインタラクション

画像の表示・非表示

ユーザが左手の5本指のジェスチャを手のひらを自分に向けて行うことでプライベート画面上に画像を表示、あるいは非表示できる（図 3.2）。



図 3.2: 画像の表示・非表示

画像の切り替え

ユーザが右手の1本指のポインティングジェスチャでプライベート画面上の画像を突き抜き、スライドすることで画像切り替えることができる（図 3.3）。



図 3.3: 画像の切り替え

画像の拡大縮小

ユーザが右手の2本指のジェスチャでプライベート画面上の画像を突き抜き、ピンチ操作することで画像を拡大縮小できる（図 3.4）。



図 3.4: 画像の拡大縮小

画像の回転

ユーザが両手の1本指のポインティングジェスチャでプライベート画面上の画像を突き抜き、回す動作することで画像を回転させることができる（図 3.5）。



図 3.5: 画像の回転

3.2.2 パブリック画面上の画像に対するインタラクション

ユーザは5本指の外向きのジェスチャを行うことでパブリック画面に対してインタラクションを開始することができる。そしてパブリック画面 GUI は図 3.6 のようになっており、その上の画像に対して以下の2つのインタラクションをカーソルを見ながら行うことができる。



図 3.6: パブリック画面の GUI

画像の配置変更

ユーザがパブリック画面上の画像にカーソルを合わせ、右手の1本指の押し込むジェスチャで画像をドラッグすることができる(図 3.7 の左)。

画像の回転, 拡大縮小

ユーザがパブリック画面上の画像にカーソルを合わせ、両手の1本指の押し込んで動かすジェスチャで画像を拡大縮小及び回転させることができる。(図 3.7 の右)。



図 3.7: 画像の位置変更及び回転, 拡大縮小

3.2.3 プライベート及びパブリック画面間での情報共有

ユーザはハンドジェスチャで自分のプライベート画面上の画像をパブリック画面に提示したり、パブリック画面上の画像を自分のプライベート画面に取得することができる。

画像の提示

ユーザがプライベート画面上の画像に対して右手の1本のポインティングジェスチャを3秒間以上行うことで画像をパブリック画面上に提示することができる(図3.8の左)。送信された画像がパブリック画面上で表示され、他のユーザからでもインタラクション可能な状態になる。

画像の取得

ユーザがパブリック画面上の画像にカーソルを合わせ、右手の1本指の押し込むジェスチャを3秒間以上行うことでその画像をプライベートプライベート端末内に取得できる。(図3.8の右)。そして、取得した画像がプライベート画面上でインタラクション可能な状態になる。



図 3.8: 画像の共有と取得

第4章 実装

4.1 開発環境及びシステム構成

本システムではユーザ側のプライベート端末と外部のパブリック端末，及び無線通信可能なネットワーク環境から構成されている（図 4.1），

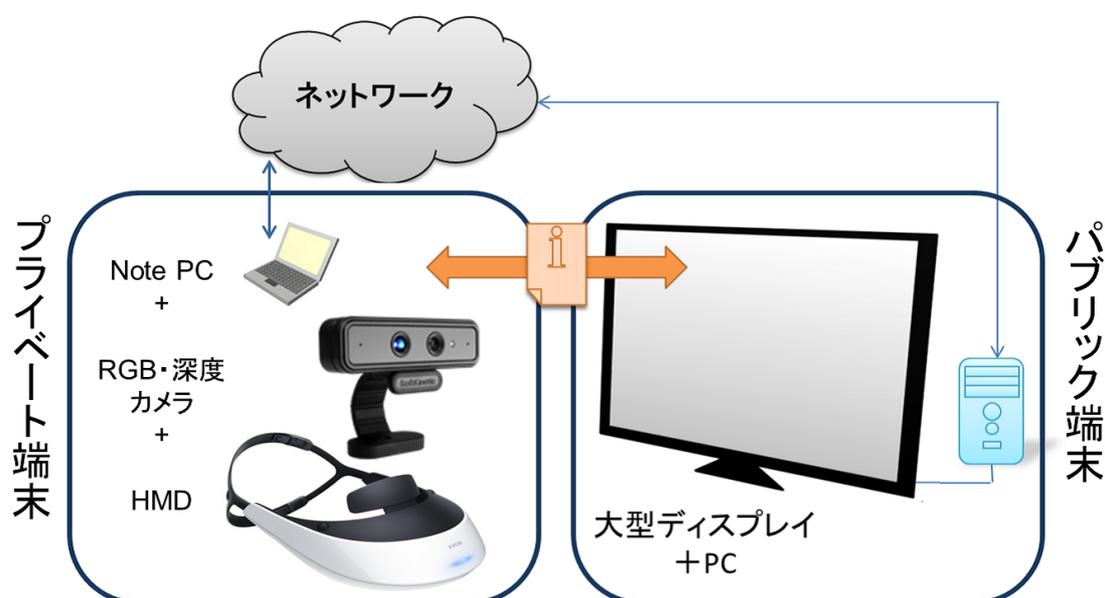


図 4.1: システム構成

プライベート端末はビデオ透過型 HMD 及びノート PC から成っており，ビデオ透過型 HMD は非透過型 HMD（SONY 社の HMZ-T2¹）に RGB・深度カメラ（SoftKinetic 社の DepthSense 325²）を取り付けることで実現している．ノート PC（Core i5 2.70GHz CPU, 4GB RAM）は無線インターネット通信及び HMD 接続用の HDMI インタフェースを搭載したものをを用いた．パブリック端末として PC（Core i5 3.20GHz CPU, 4GB RAM）に接続した 50 インチの大型液晶大画面を用いた．本研究で用いた 2 つの PC の OS は 64 ビット Windows³ でした．以降，実

¹<http://www.sony.jp/hmd/products/HMZ-T2/>

²<http://www.softkinetic.com/Store/tabid/579/ProductID/6/language/en-US/Default.aspx>

³<http://windows.microsoft.com/>

装の説明ではユーザ側であるプライベート端末側を HMD 側、パブリック端末側を大画面側と呼ぶ。

HMD 側のアプリケーションは C++ 言語を用いて開発した。大画面側のアプリケーションは Java 言語を使用して開発を行った。RGB・深度カメラからの画像情報や深度情報を取得するために SoftKinetic 社の DepthSenseSDK⁴ を利用している。そして、画像処理用のライブラリー OpenCV⁵ を用いてハンドジェスチャの認識を行った。また、大画面の特定、識別には AR マーカ及び ARToolKit⁶ を用いている。

4.2 通信方式

本研究では HMD 側と大画面側との間に無線ネットワークを通じてジェスチャ情報及びその他の情報を送受信する。通信方式として UDP 及び TCP プロトコルを用いており、通信するメッセージの内容によって使い分けている (図 4.2)。

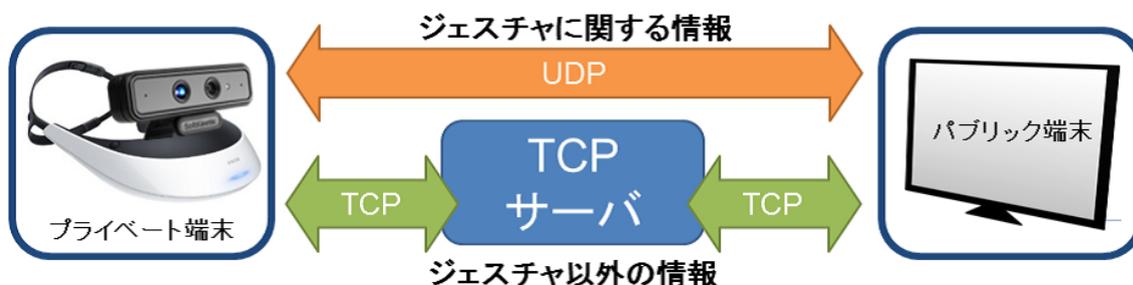


図 4.2: 通信方式

HMD 側からのジェスチャの種類やジェスチャの座標に関する情報がリアルタイムで大画面側に送信される必要があるため UDP プロトコルを用いた。そのために大画面側を UDP サーバとして常に HMD 側からの通信を待ち受け状態にし、HMD 側を UDP クライアントとしている。UDP プロトコルではパケットロスが生じるが、それによるエラーをアプリケーション側で補正している。一方で、端末間のファイル通信には信頼性の高い TCP プロトコルを用いて送受信される。その際に、複数端末の連携を想定して HMD 及び大画面を TCP クライアントとして 1 つのサーバを通じて通信させる。TCP サーバがクライアントからのメッセージを分析して、適切な端末間でメッセージを通信させる。例えば、TCP サーバでは大画面に関する情報を管理しており、HMD 側に選択された画面の情報を送信したり、HMD と大画面の間で画像情報を通信したりする。

⁴<http://www.softkinetic.com/en-us/support/download.aspx>

⁵<http://opencv.org/>

⁶<http://www.hitl.washington.edu/artoolkit/>

4.3 大画面の識別及び位置関係

4.3.1 大画面の識別

本研究では大画面の識別に AR マーカを用いる (図 4.3). 具体的には, 大画面にそれぞれ固有 ID を持つ AR マーカを貼り付け, それを HMD の RGB カメラで探索し, 特定することで画面を識別する. AR マーカの ID と大画面情報 (大画面の ID, IP アドレス, サイズ, マーカとの相対関係) がそれぞれ関連付けられ, サーバで管理される.

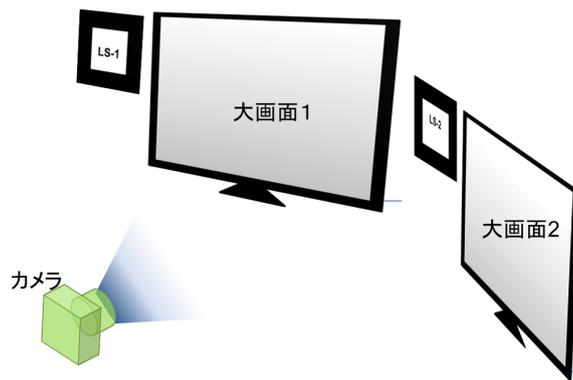


図 4.3: 大画面の特定

4.3.2 大画面の位置及び向き

大画面と HMD の位置関係を求めることで, インタラクションの際の大画面の動的選択を可能にする. そのために, 大画面を特定した後, そのマーカに対して ARToolKit の `arGetTransMat()` メソッドでマーカ座標系 (X_m, Y_m, Z_m) からカメラ座標系 (X_c, Y_c, Z_c) への変換行列 T を求める (図 4.4). これを数式で表したのが式 4.1 で, 簡略化したものが式 4.2 である. 変換行列 T の r 要素は回転成分であり, t 要素は並進成分である.

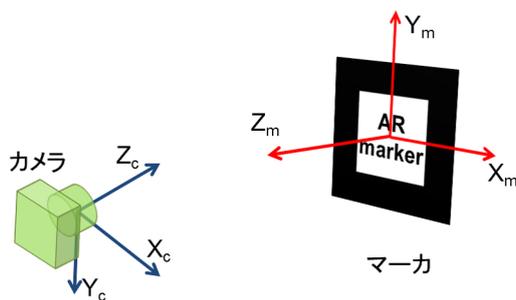


図 4.4: カメラ座標系とマーカ座標系

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \quad (4.1)$$

$$C = T \cdot M \quad (4.2)$$

そして、識別された画面の解像度と HMD のカメラ視野（ジェスチャ認識可能な領域）を対応させ、大画面上にユーザのハンドジェスチャのフィードバックとなるカーソルを表示させる（図 4.5）。

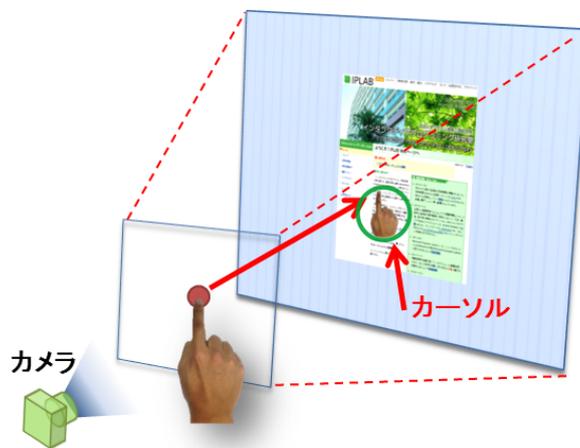


図 4.5: 大画面上でのハンドジェスチャのフィードバック

4.4 アプリケーションの実装

HMD 側では深度カメラを用いてハンドジェスチャを認識し、画面上の画像をアフィン変換で拡大縮小、回転などを行うアプリケーションを実装した。その際に画像処理ライブラリーである OpenCV を用いた。また、ユーザが大画面に対して行ったジェスチャの位置や種類などの情報を大画面側に送信するネットワーク機能も実装した。

大画面側では、画像情報データを提示、移動、回転、拡大縮小を行うアプリケーションを Java の Swing を用いて Eclipse 3.7.1 で作成した。JFrame 上に写真を読み込み Graphics2D のアフィン関数を用いて画像の平行移動、回転、拡大縮小を行った。そして、HMD 側からの通信を待ち受けし、ハンドジェスチャの情報をリアルタイムで取得し、コマンドを実行するネットワーク機能を実装した。

4.5 ハンドジェスチャの認識

ハンドジェスチャ認識は次の6つの処理を順次に繰り返し行うことで実現している。

1. 両手の領域を検出する。
2. 検出された各領域が左手か右手か判別する。
3. 手の軸を求める。
4. 手の各指（指先）を検出する。
5. 手の向きを判別する。
6. ハンドジェスチャを分類する。

そして、これらの処理を行うメソッドや変数を階層的なクラスを独自に作成して管理している。例えば、ハンドジェスチャの種類を管理する `GestureType` クラスや、ハンドジェスチャの情報を管理する `Gesture` クラス、手の情報を管理する `Hand` クラス、指の情報を管理する `Finger` クラス、点集合を管理する `Points` クラスなどがある。以下に各処理を説明する。

4.5.1 両手の領域の検出

本研究の構成では、HMD から手までの距離はほぼ腕の長さに等しいという特徴があるため、深度カメラを用いて一定距離以内物体のみを検出することで両手以外の背景を認識から除去することができる。試行錯誤の結果、開発者の両手はカメラから 800[mm] 以内の距離に収まるということが分かった。この基準を基に深度カメラから各画素が被写体までの距離を [mm] 単位で表した 320 x 240 サイズの `DepthMap`（深度マップ）を取得する（図 4.6 の左）。また、同時に各画素の深度情報の信頼度で構成された同じスケールサイズの `ConfidenceMap`（信頼度マップ）を取得する（図 4.6 の右）。

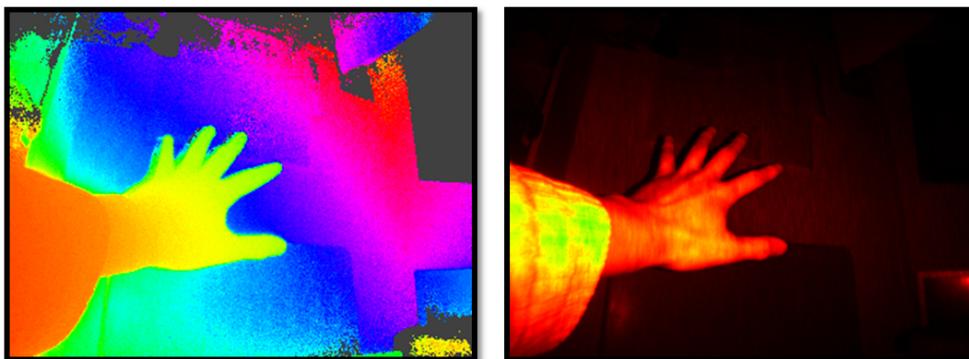


図 4.6: `DepthMap` 及び `ConfidenceMap`

ConfidenceMap は DepthMap に比べてノイズが少ないため手の領域を求めるマスク画像として使う (図 4.7 の左). そしてそのマスク画像を DepthMap に適用することでよりノイズが少ない深度マップを取得できる (図 4.7 の右).

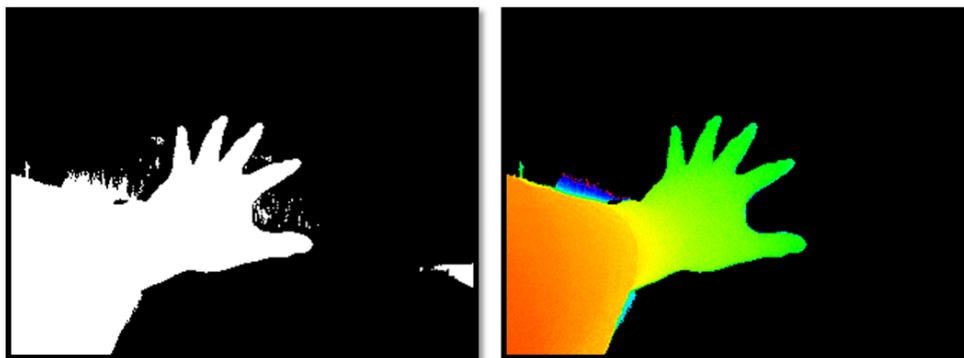


図 4.7: 手のマスク画像及び手の深度マップ

4.5.2 各領域の左手か右手かの判別

カメラがユーザの目の位置に設置されている本研究の構成では基本的に右手領域の重心は画面右側, 左手領域の重心は画面左側に存在するという特徴がある. この特徴を使い, まず検出された各領域 (図 4.8 の左) の重心の位置を求め, それをもとに左手か右手かを判別する (図 4.8). 手全体の領域の重心を OpenCV の Moments() メソッドを用いて領域の輪郭情報から求める.

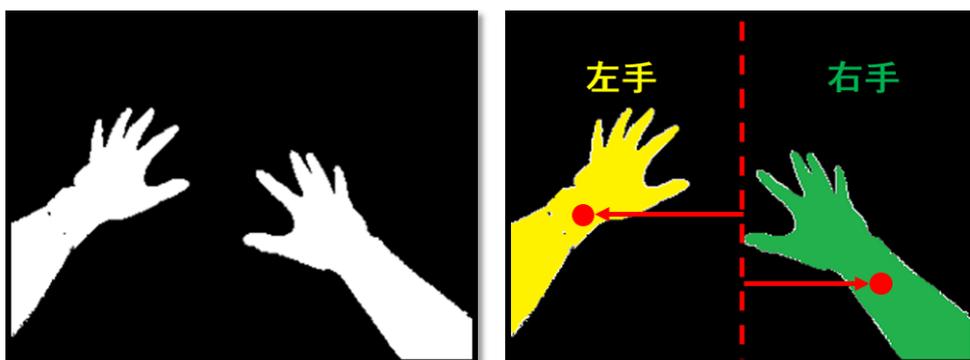


図 4.8: 手の左右の判別

4.5.3 手の軸の検出

本研究において手全体の重心（図 4.9 の緑色の丸）と手の先端（図 4.9 の赤色の丸）を結ぶベクトルを手の軸と呼んでおり，手のひらの向きを判別するための重要な要素である．指や手の向きを手のひらの中心点を用いて検出できるが，本研究では手の先端に注目した手の軸を用いている．その理由は手の先端のほうがより認識しやすく，インタラクションの際に安定して追跡できると考えたからである．

手の先端は手の領域に楕円形（図 4.9 の緑色の曲線）を適用しその鋭い先端を選択することで求める．楕円形には鋭い先端が 2 つ存在するが，画面中央に近いほうが手の先端であると判断する．

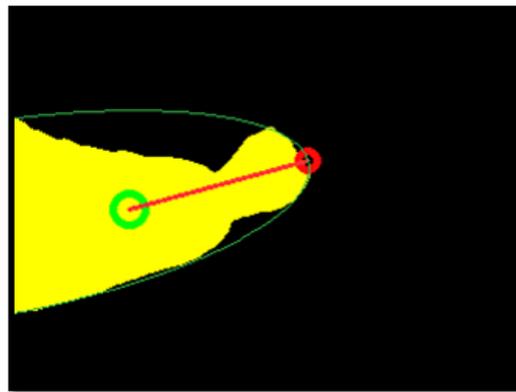


図 4.9: 手の先端及び手の軸の検出

4.5.4 指先の検出

各手の輪郭から次の条件を満たす形状の部分指先だと判断する．

輪郭の各点 P_i や，同じ輪郭上の k 前の点 P_{i-k} と k 後の点 P_{i+k} の 3 つの点について，

- 3 つの点のなす角度 $\angle P_{i-k}P_iP_{i+k}$ が 30 度以下であること
- 点 P_i が手の領域の凸閉包上の点であること

1 つ目の条件により指のような凸型の部分が検出される (4.10-a)．2 つ目の条件により指先と同じような鋭い角度を持つ指同士の間を谷を検出から除外する．その結果，指先付近の点だけが検出される．そして，本研究では k を点 P_i の深度情報をもとに動的に変化するようにした．具体的には，深度が 100[mm] のとき k が 55 で，800[mm] のとき k が 5 になるように次のような線形関数を作った (4.1)．そして， k の最大値は 55，最小値は 5 に設定した．

$$k = (870 - \text{点 } P \text{ の深度})/14 \quad (4.3)$$

この2つの条件で検出された点は1つの指に対して複数存在するため Point 型の点の集合 Points クラスを作成して1つの指先を近くに検出された複数の指先の候補点の集合として管理する (4.10-b)。そして、その点の集合の中心点をその指先だと判断している (4.10-c)。最後に、この2次元座標を用いて深度マップから深度情報を取得し、指先の3次元座標を取得する。

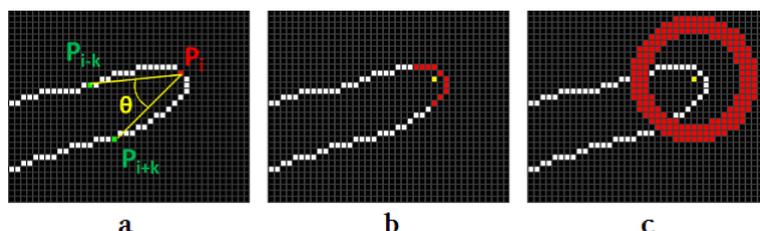


図 4.10: 指先検出

4.5.5 手のひらの向きの判別

親指の検出

手のひらの方向を認識するために親指の指先の位置を用いている。親指の指先は他の指先から最も離れている特徴を利用して親指を求める。例えば以下の図 4.11 の左では他の指先と最も離れている指先が親指と認識されて紫色の円で示されている。

手の軸と親指の位置関係

手の左右がわかっているときは親指が手の軸に対してどこにあるかで手のひらの向きを判別できる。例えば手が左手だとわかっている場合、親指の位置が手の軸から見て左側にあったら手のひらがユーザに向いているということがわかる (図 4.11 の右)。

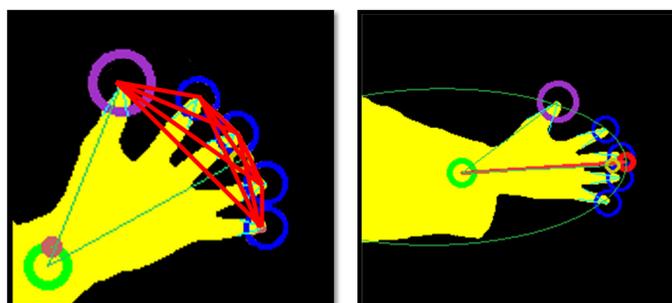


図 4.11: 親指の検出と手の左右の判別

4.5.6 ジェスチャ判別

両手やその指を求めた後，検出された手の左右や指の数でジェスチャの種類を判別する．例えば，認識された手の数が1で，指の数が1であれば片手で指差しているジェスチャだと判断させる．片手のみのジェスチャでは，指の数が0から5まで6種類（図4.12）あり，その中で親指の位置が特定しやすい5本指のジェスチャ（図4.12 - F, H）及び3本指のジェスチャ（図4.12 - D, G）は手の向きも認識できるため合計8種類のハンドジェスチャを認識できる．左右の手で合計16種類になる．両手同時ののハンドジェスチャは，手それぞれ8種類とし，合わせて64通りのハンドジェスチャを認識できる．つまり単純計算で80通りの静的ハンドジェスチャを認識できる．

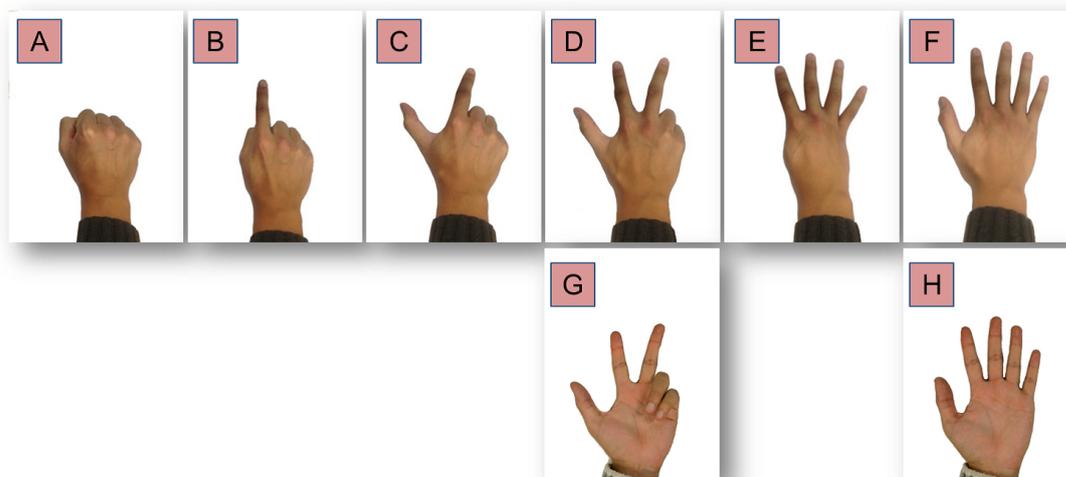


図 4.12: 片手のハンドジェスチャ

4.6 指で突き抜く動作の判定及びオクルージョンの再現

4.6.1 突き抜く動作の判定

ここではハンドジェスチャの認識で得られた指先の3次元情報を用いる．具体的には，各指先の2次元座標が画像が表示されている範囲内であって，深度情報が400[mm]より以上になるとHMD上の画像を突き抜く動作が行われたと判定する．つまり，HMD上の画面がユーザから400[mm]の距離にあるかのように設定されている．

4.6.2 オクルージョンの再現

本システムでは、カメラから 400[mm] 以内の距離の深度を撮ったマスク画像（図 4.13 の左）を用いてインタラクションの際にユーザの手と HMD 上の画像がユーザからの奥行きによって隠れて見えるように再現した。例えば、図では画像を突き抜けた手の部分が画像で隠れていることが分かる（図 4.13 の右）。

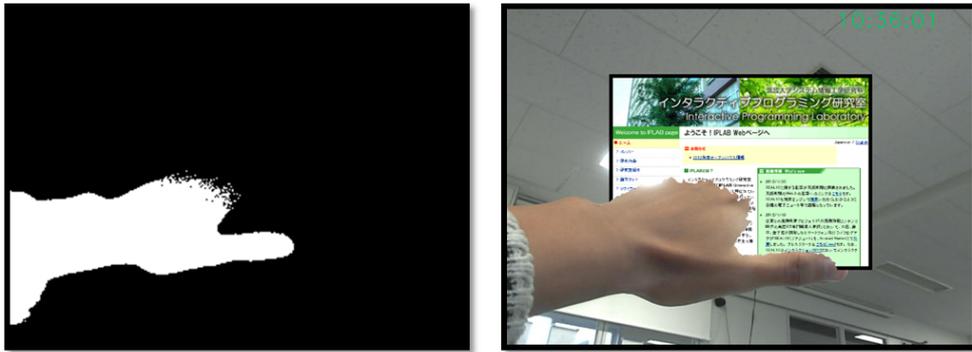


図 4.13: オクルージョンの再現

第5章 関連研究

5.1 HMD インタラクションに関する研究

本研究ではHMDにおけるハンドジェスチャインタラクションシステムを開発しており、ウェアブルコンピュータインタラクション及びARインタラクション、特にその中でハンドジェスチャによるインタラクションシステムに関する研究分野に関連している。この分野で従来より様々な研究が行われてきたが、以下に本研究がそれらと異なる点について述べる。

山本ら [20] が手や足に取り付けたセンサによってHMDとインタラクションを行うシステムを提案した。Parkら [17] もユーザの視線を検出することでHMD内の情報を選択できるシステムを提案した。本研究はHMDとインタラクションを行うシステムとしてこれらの研究と関連するが、RGBカメラと深度カメラを用いて多種類のジェスチャを認識することでより豊かなインタラクションを可能にしているところが異なる。

また、Kimら [12] のDigitsというユーザが手首に付けることでハンドジェスチャによる様々なインタラクションが可能なるシステムを提案している。このシステムは手首の内側に赤外線LEDと赤外線カメラを取り付けることで各指の位置をモデル化し、手の姿勢を認識できるシステムである。本研究ではユーザが手に何も付ける必要がないため、より負担が少ないハンドジェスチャを実現している点でこの研究と異なる。

蔵田ら [13] はHMDに取り付けたRGBカメラを用いてハンドジェスチャを認識できるシステムを提案し、その応用を示した。このシステムでは手の色情報を用いてハンドジェスチャ認識し、ポインティングやフリック操作を実現している。また、認識における環境照明による影響を減らすための独自のアルゴリズムを開発している。本研究が深度カメラを用いて環境照明に影響を受けない、よりロバストなハンドジェスチャ認識方法を実装することで、より豊かなインタラクションを可能にしているところがこの研究と異なる。

加茂ら [10] が拡張現実感を用いた仮想タッチパネルインタフェースを提案し、金子ら [11] がそれを用いた協調作業支援システムを提案した。これらのシステムはHMD上で出現させた仮想仮面をハンドジェスチャによってインタラクションを行うことができるシステムとして本研究と関連するが、本研究がHMD上で見える画面だけではなく大画面インタフェースもハンドジェスチャで簡単に切り替えてインタラクションを行うことができるところが異なる。

Colaoら [5] がMimeというHMD用の低電力ジェスチャインタラクションシステムを提案した。3つのフォトダイオードを用いてユーザの手の位置及び平均深度を求め、RGBカメラで手の詳しい姿勢を認識することで低電力を実現している。本研究は将来的に深度カメラがより小型で低電力になると想定し、深度カメラを用いてより多くのハンドジェスチャを高精度で認識することに重点をおいているところが異なる。

5.2 大画面インタラクションに関する研究

本研究の先行研究 [6][7][8] としてペンダントように装着した携帯端末の RGB カメラを用いてハンドジェスチャを認識し、大画面とインタラクションを行うことができるシステムがある。今回の研究では深度カメラを用いることで環境照明の影響を受けないよりロバストなジェスチャ認識システムを実現できた。また、HMD を用いることで個人デバイス内の情報をプライベート画面で確認できるようにした。

Argyros ら [1] は、カメラを用いてユーザの 2 次元静的ジェスチャと 3 次元静的かつ動的ハンドジェスチャを認識することでコンピュータのマウスの代わりに使うことができるシステムを提案した。Shi ら [18] は、複数のカメラを用いて手の 3 次元ジェスチャを認識し、画面を見ながら 3 次元バーチャルオブジェクトを操作できるシステムを提案した。Cheng ら [4] は大画面とインタラクションを行うための仮想タッチスクリーンシステムを提案し、大画面に取り付けたカメラでユーザの目の位置を検知することでユーザがタッチスクリーンを操作しているような体験が得られるようにした。Nakamura ら [15] は指に赤外線 LED を取り付けることでポインティングジェスチャを認識し、大画面上でのインタラクション手法を提案した。Lee ら [14] は深度カメラを用いて手や手のひら、指を認識することで大画面とインタラクションできるハンドジェスチャ手法を提案した。これらの研究は RGB カメラや深度カメラを用いてハンドジェスチャを認識し、大画面上でインタラクションを行うところが本研究と類似しているが、ジェスチャを認識するためのカメラが大画面側に固定されており、インタラクションを行う際にユーザの動きを制限してしまう問題があった。本研究は HMD に取り付けられた深度カメラでジェスチャを認識することでこの問題を解決しているところやより多種類のハンドジェスチャを認識できるところが異なる。

金ら [9] は、タッチパネル付きモバイル端末を用いてパブリック壁面大画面とインタラクションを行い、大画面とモバイル端末間で情報共有できるシステムを提案した。また、Bragdon ら [3] は、室内に設置した深度カメラで大画面をジェスチャで操作できたり、手元のモバイル端末で大画面と情報交換を行ったりできるシステムを提案した。Borrington ら [2] はスマートフォンのカメラ越しに大画面とインタラクションを行うシステムを提案した。本研究は、インタラクションの際にユーザが手に携帯端末を持つ必要がなく、ハンドジェスチャのみで様々なインタラクションや情報交換が可能になっているところがこれらの研究と異なる。

Vogel ら [19] が壁面大画面上でポインティングやクリック操作ができるハンドジェスチャインタラクションを提案した。本研究はより多種類のハンドジェスチャを認識できるところが異なる。

Nancel ら [16] は壁面大画面においてマルチタッチ可能なモバイル端末を用いたインタラクションとハンドジェスチャによるインタラクションを比較して、単純な線型的なタッチジェスチャはモバイル端末のほうが効率的という結果を示している。本研究は、未来環境において従来のモバイル端末は HMD に統合され、人々が手に端末を持たずにインタラクションできるようになると想定しており、ハンドジェスチャによる豊かなインタラクションを重視しているところが異なる。

第6章 課題と今後の発展

6.1 課題

本研究で実装行ったシステムを試用した結果いくつかの問題点を発見した。本研究でビデオ透過型 HMD を用いたためユーザの視界として HMD に取り付けられたカメラの視界を用いており、ユーザの手を認識できる範囲も限られていた。ユーザがジェスチャを行うとき手を高く維持しなければならず、インタラクションの際に手にかかる負担が大きいという問題があった。将来的に視界が人間の目の視界と同等な広さの HMD や RGB・深度カメラが実現されるときこの問題は解決されると考えられる。そして、システムを客観的に評価し、ユーザにとってより使いやすいハンドジェスチャやインタラクションを検討していく必要があると考えられる。

ジェスチャ認識に関しては、限られているハンドジェスチャの場合のみ手の向きを検出できるという問題があった。手の向きを親指と手の軸を用いて検出しているため、親指が見えないなど認識が困難なジェスチャの場合、手の向きは認識できなかった。この問題は、手のひらの形状やその中心点の位置を正確に検出し、各指先との距離を基に親指を正確に検出することで解決できると考えられる。そのため、今後の研究で手の平を正確に認識できるアルゴリズムを開発にも力を入れたいと考えている。そして、ジェスチャ認識の精度を定量的に評価し、既存のシステムと比較を行う必要があると考えられる。

また、今回実装したシステムはビデオ透過型 HMD を用いて実装しているが、より実用的なシステムにするために、今後ハーフミラーによる透過型（光学透過型）HMD を用いて実装を行う必要と考えられる。

6.2 今後の発展

今回の研究では画像データのみ扱っていたため、今度画像以外の多種類のデータ（例えば、音声データ、ビデオ、PDF、プレゼン資料など）にも対応させることが考えられる。そして、より多種類で細かなハンドジェスチャを認識できるようにし、ユーザそれぞれが自分に合った直感的なハンドジェスチャでインタラクションできるシステムにすることが考えられる。また、今回は大画面の特定に AR マーカを用いたが、今後マーカレス AR 技術を用いてユーザが大画面を見るだけで大画面を選択できるシステムを開発することが考えられる。

ハードウェアの面では、今後 Google Glass のような光学透過型 HMD に対応させ、誰もが自由に利用できるシステムにすることが考えられる。

第7章 結論

本研究では，将来複数大画面や HMD が混在した環境が実現されると想定し，そのような環境におけるプライベート及びパブリック画面の操作に適したハンドジェスチャインタラクションシステムを提案した．そして，そのプロトタイプシステムとして HMD と大画面を使った画像共有アプリケーションを実装した．本システムでは，ユーザが装着した HMD の RGB・深度カメラを用いてハンドジェスチャを認識することでユーザが従来のシステムのように大画面の前で拘束されることなく自由にインタラクションを行えることが確認できた．また，複数大画面環境において見ている方向のみで大画面を動的に切り替えることができた．そして，大画面インタラクションの際に HMD を個人識別デバイスとして使うことで正確にユーザ識別し，大画面とユーザの間での情報共有を行うことができた．

謝辞

本研究を遂行するにあたり，ご指導頂きました田中二郎教授に心より感謝申し上げます。そして，三末和男准教授，高橋伸准教授，志築文太郎准教授，嵯峨智准教授，Simona Vasilache 助教には日々のゼミや研究室での日常生活の中で様々な貴重なご意見やご指導を頂きました。本当にありがとうございました。また，学生生活において日頃からお世話になっていた先輩方や同期の方々，後輩達に感謝申し上げます。私に勉強できるこのような貴重な機会を与えてくれた筑波大学や日本政府にも感謝申し上げます。

最後に，精神面でも金銭面でもいつも支えてくれ，応援してくれた両親や家族，そして留学生活の中で様々な面においてお世話になったつくば市内に住む同じ国の留学生たちやその家族の方々にこの場を借りて心から感謝申し上げます。

ありがとうございました。

参考文献

- [1] Antonis A. Argyros and Manolis I. A. Lourakis. Vision-based interpretation of Hand Gestures for Remote Control of a Computer Mouse. *Proceedings of the HCI '06 workshop*, pp. 40-51, 2006.
- [2] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: mobile interaction through video. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, pp. 2287-2296, 2010.
- [3] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. Code space: touch + air gesture hybrid interactions for supporting developer meetings. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pp. 212-221, 2011.
- [4] Kelvin Cheng and Masahiro Takatsuka. Estimating virtual touchscreen for fingertip interaction with large displays. *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments (OZCHI '06)*, pp. 397-400, 2006.
- [5] Andrea Colao, Ahmed Kirmani, Hye Soo Yang, Nan-Wei Gong, Chris Schmandt, and Vivek K. Goyal. Mime: compact, low power 3D gesture sensing for interaction with head mounted displays. *Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST '13)*, pp. 227-236, 2013.
- [6] ダワースレン エンクバト. 装着型携帯端末を用いた大画面用ジェスチャインタラクション. 卒業論文, 筑波大学, 2011.
- [7] Davaasuren Enkhbat, 田中二郎. 装着型携帯端末による大画面ジェスチャインタラクションの提案. マルチメディア, 分散, 協調とモバイル (DICOMO2012) シンポジウム, pp. 634-645, 2012.
- [8] Enkhbat Davaasuren and Jiro Tanaka. MOBAJES : Multi-User Gesture Interaction System with Wearable Mobile Device. *Proceedings of 15th International Conference on Human-Computer Interaction (HCI International 2013)*, Human-Computer Interaction: Interaction Modalities and Techniques (Part IV), pp.196-204, 2013.
- [9] 金春明. 携帯デバイスを入力装置として用いた公共空間の大画面インタラクション手法. 修士論文, 筑波大学, 2007.

- [10] 加茂浩之, 田中二郎. ウェアラブル拡張現実感による情報端末の仮想化. マルチメディア, 分散, 協調とモバイル (DICOMO2011) シンポジウム, pp.1223-1233, 2011.
- [11] 金子将大, 田中二郎. 拡張現実感を利用した仮想的なディスプレイによる作業の支援. マルチメディア, 分散, 協調とモバイル (DICOMO2012) シンポジウム, pp. 433-441, 2012.
- [12] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor. *Proceedings of the 25th annual ACM symposium on User interface software and technology (UIST '12)*, pp. 167-176, 2012.
- [13] 蔵田武志, 興梠正克, 加藤丈和, 大隈隆史, 坂上勝彦. ハンドマウスとその応用: 色情報と輪郭情報に基づく手の検出と追跡. 映像情報メディア学会技術報告, Vol. 25, No. 85, pp. 47-52, 2001.
- [14] Unseok Lee and Jiro Tanaka. Finger Controller: Natural User Interaction using Finger Gestures. *Proceedings of 15th International Conference on Human-Computer Interaction (HCI International 2013)*, Human-Computer Interaction: Interaction Modalities and Techniques (Part IV), pp.281-290, 2013.
- [15] 中村卓. 大画面環境におけるポインティングハンドジェスチャによるインタラクション手法の研究. 博士論文, 筑波大学, 2012.
- [16] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. Mid-air pan-and-zoom on wall-sized displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, pp. 177-186, 2011.
- [17] Hyung Min Park, Seok Han Lee, and Jong Soo Choi. Wearable augmented reality system using gaze interaction. *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pp. 175-176, 2008.
- [18] Junhao Shi, Mingmin Zhang, and Zhiheng Pan. A real-time bimanual 3D interaction method based on bare-hand tracking. *Proceedings of the 19th ACM international conference on Multimedia*, pp. 1073-1076, 2011.
- [19] Daniel Vogel and Ravin Balakrishnan. Distant freehand pointing and clicking on very large, high resolution displays. *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*, pp. 33-42, 2005.
- [20] 山本哲也, 庄司武, 寺田努, 塚本昌彦. ウェアラブルコンピューティングのための手足を使った状況依存コマンド入力手法. インタラクション 2011 論文集, pp.99-106, 2011.