平成23年度

筑波大学情報学群情報科学類

卒業研究論文

題目
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

主専攻　　情報システム主専攻

著者　　Davaasuren Enkhbat

指導教員　田中二郎　志築文太郎　高橋伸　三末和男

要　旨

学校や職場など様々な場所で協調作業を行う際に，大画面を用いて作業する場面が今後ますます増えると考えられる．そのとき，複数人で複数の大画面とインタラクションを行い，情報共有できるインタラクション手法が必要になると考えられる．従来のインタラクション手法であるマウスやタッチパネルによる操作では，複数大画面と情報共有を行うことが難しい．

そこで我々は複数大画面を使った協調作業において直感的に情報共有を行うことができるインタラクションシステムとして装着型携帯端末を用いたハンドジェスチャインタラクションシステムを提案する．本手法では，ユーザが背面にカメラが付いている携帯端末をカメラがユーザの前方を向く状態でペンダントのように身につけ，その前で両手によるジェスチャを行うことで大画面とインタラクションを行い，情報共有ができる．具体的には情報を大画面上で操作することや，大画面上から取得すること，大画面上に提示することなどが可能になっている．

目 次

第 1 章　はじめに **1**
1.1 背景 1
1.1.1 大画面の閲覧性と表現力 1
1.1.2 大画面上の情報の携帯性 1
1.1.3 従来の大画面上での情報共有の問題点 2
情報提示 2
情報取得 2
1.1.4 従来の大画面インタラクション手法の問題点 2
マウス 2
タッチパネル 2
音声 3
1.2 本研究の目的 3
1.3 本研究のアプローチ 3
1.4 本論文の構成 3

第 2 章　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 **4**
2.1 想定する未来環境 4
2.2 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とは 4
直感的な操作性が得られる 5
情報共有ができる 5
コンピュータを意識しない 5
2.3 想定する利用シナリオ 5

第 3 章　プロトタイプシステム「MobaJes」 **7**
3.1 システムの概要 7
3.2 インタラクション 7
3.2.1 基本ジェスチャ 7
3.2.2 大画面上でのインタラクション 8

第 4 章　システムの実装 **10**
4.1 開発環境と言語 10
4.2 システム構成 11

4.3 通信方式 11
4.4 大画面側アプリケーション 13
4.4.1 GUI の実装 13
4.5 携帯端末側アプリケーション 13
4.5.1 ジェスチャを管理するクラスの作成 13
4.5.2 ジェスチャ認識 15
肌色領域の検出 15
手とその重心点の検出 16
指先の検出 17
ジェスチャ判別 18
4.5.3 大画面との接続 18
ユーザが向いている方向の検出 18
大画面の選択 19

第 5 章 関連研究 **21**
5.1 携帯端末を用いた大画面インタラクション 21
5.2 ハンドジェスチャを用いた大画面インタラクション 21
5.3 大画面との情報共有 22
5.4 装着可能なジェスチャインタフェース 22

第 6 章 議論 **24**
6.1 試用から得られたシステムの利点と問題点 24
6.2 今後の展望 25

第 7 章 結論 **26**

謝辞 **27**

参考文献 **28**

図 目 次

2.1 提案システムのイメージ 4

3.1 指す 7
3.2 掴む 7
3.3 放す 7
3.4 配置の変更 8
3.5 表示サイズや傾きの変更 8
3.6 取得 8
3.7 提示 8

4.1 携帯端末を装着している様子 10
4.2 システム構成 11
4.3 通信方式 11
4.4 大画面側アプリケーションの GUI 13
4.5 ジェスチャを管理するクラスの関連 14
4.6 肌色領域に検出.a) 元画像, b) 肌色領域の抽出, c) ノイズ除去後の画像. 16
4.7 手とその重心点の検出.a) 輪郭の抽出, b) 面積によるフィルタリング, c) 領域の重心点. 17
4.8 指先検出.a) 凸型部分の検出, b) 指先の候補点の集合, c) 指先判定. 18
4.9 携帯端末の X,Y,Z 軸 19
4.10 大画面の選択（部屋の真上から見た場合） 20

第1章　はじめに

1.1　背景

1.1.1　大画面の閲覧性と表現力

協調作業においては協調者たちの間で情報共有を行うことがとても重要なことである．情報共有によって全体やお互いの状況が把握しやすくなり，作業がスムーズに行われるからである．複数人で情報共有を行う際に最も多く用いられる方法はプロジェクタや大型ディスプレイなど大画面を用いたプレゼンテーションやミーティングである．実際に，学校や職場では以前から日常的に行われてきた．それは，人間が外部から得られる情報の 83 %が視覚によるものだと考えられており [1]，情報共有に置いて閲覧性が高い大画面を用いた方が情報を与える側と受け取る側の両方にとって，最も効率的だからだと考えられる．

また，大画面はホワイトボードや掲示板ではできない動画やアニメーションなどを用いて，情報をわかりやすく提示できるため，高い表現力を持っていると考えられる．

そして，大画面は今後も情報共有の場として重要度がますます高まり，特に学校や職場における協調作業の際に複数の大画面を使って作業する場面が多くなると予想される．

1.1.2　大画面上の情報の携帯性

大画面には高い閲覧性や表現力の反面，大画面上での情報の携帯性が乏しいという問題がある．ここで，大画面上での情報の携帯性とは，情報が複数大画面間で，あるいは大画面とユーザの間で自由に移動できることや操作が可能であることを意味している．

つまり今後，協調作業において複数の大画面が使われるようになったとき，大画面上での情報の自由な操作や，自由な情報共有といった情報の携帯性が重要にになると考えられる．しかし，従来の大画面情報共有方法や大画面インタラクション手法では情報を自由に操作したり，共有したりすることが困難である．

そのため，従来の大画面情報共有方法や大画面インタラクション手法の問題点を解決した，新しい情報共有方法やインタラクション手法が必要である．

1.1.3　従来の大画面上での情報共有の問題点

情報提示

大画面上で情報を提示する際に 1 つの情報源から一方的に情報提示する場合が多い．例えば，会議やミーティングの際に，発表者が大画面に自分のパソコンを接続して情報を提示する．そのため，大画面が発表者 1 人の支配下におかれ，協調者たちが情報を提示する機会が失われてしまう．そして，大画面上での情報の流れが一方的になり，議論がうまく進まないという問題がある．また，情報提示はパソコンなどを通して行うため，コンピュータやその操作方法を意識する必要があり，ユーザにとって負担になると考えられる．

情報取得

大画面で共有されている情報から必要な情報を取得する従来の方法として，メモを取る，写真に写すなど方法がある．しかし，そのように取得した情報はその情報が持つ本来の性質が失われる問題がある．例えば，大画面上のテキストファイルを写真として取得した場合，テキストファイルに戻すことが難しく，内容からキーワードを検索するなど本来の簡単にできる操作ができなくなる．また，取得した情報に余計な情報が含まれていたり，重要な部分が抜けていたりする問題も考えられる．そして，情報を再共有することも難しくなり，情報の携帯性が落ちると考えられる．

1.1.4　従来の大画面インタラクション手法の問題点

マウス

現在使われている最も古いインタラクション手法の 1 つで，ほとんどのデスクトップ PC で利用されているインタフェースである．GUI に適して作られており，水平面の上で動かすことで，カーソルを移動させ，画面上の情報を操作するものである．そのため，大画面とインタラクションする際は，手を動かす量が多くなり，ユーザに負担を与えてしまう，手を動かす面と大画面が垂直になり，インタフェースとして自然な対応が得られないなど問題がある．

タッチパネル

現在，スマートフォンやタブレットデバイスなどで主に使われているインタラクション手法であり，ユーザが手で画面上の情報にタッチすることで情報に対して様々な操作するものである．インタラクションを行う際に手で何かを持つ必要がなく，手を自由に動かしながら操作できるため自由度の高いインタラクション手法である．また，情報に直接触れているような感覚をユーザに与えるため，自然な対応が得られる．しかし，インタラクションを行うために手で画面に触れる必要があるため，大画面の大きさや設置場所によってインタラクションできなくなってしまうという操作の自由度の問題がある．

音声

ユーザが音声で命令を出すことでインタラクションを行う手法である．言語を使って豊富な種類の命令を出すことが可能であり文字入力や複雑なインタラクションの際に有効である．しかし，図書館など静かな公共場で使うことが難しく，また，複数人でインタラクションを行う場合や，野外など音が多くする環境で使う場合にユーザの意図しない誤動作をする問題がある．

1.2 本研究の目的

複数大画面を用いた協調作業において，従来の情報共有方法では，提示される情報が一方的で限られた情報になることや，情報取得方法によって情報やその性質が失われてしまうことなど問題点があった．また，従来のインタラクション手法では自由度の高い操作性や自然な対応が得られない，離れた場所から，あるいは複数人で大画面とインタラクションできないなど問題があった．

そこで本研究の目的として，これらの問題を解決した新しいインタラクション手法を提案し，複数大画面を使った協調作業において大画面と直感的にインタラクションを行いながら，コンピュータを意識せずに情報共有できるシステムを開発する．

1.3 本研究のアプローチ

本研究では，カメラ付き装着型携帯端末を用いたハンドジェスチャインタラクション手法を用いる．これにより，両手のジェスチャによる高い操作性が得られ，自由に情報共有できるようなる．また，ユーザが装着型携帯端末を直接操作しないため，コンピュータを意識する必要がなくなる．

1.4 本論文の構成

本論文の構成は以下の通りである．

本章では協調作業における大画面の重要性や情報の閲覧性の必要性について説明し，大画面における従来の情報共有方法やインタラクション手法の問題点を解決する筆者のアプローチや研究の目的を述べた．第 2 章では本研究で提案する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システムについて説明する．第 3 章では本研究で実装した「MobaJes」プロトタイプシステムについて説明し，そのインタラクション方法について述べる．第 4 章では「MobaJes」システムの実装について説明する．第 5 章では関連研究について述べて，本研究との相違点について説明し，研究の位置づけを明確にする．第 6 章ではシステムの利点や問題点について議論し，問題点の改善策やシステムの今後の発展に付いて議論する．そして，最後の第 7 章では結論を述べる．

第2章　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

2.1　想定する未来環境

本研究では，近未来に学校や職場など協調作業を行うあらゆる場所で複数の大画面が使用されるようになると想定している．

2.2　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とは

図 2.1: 提案システムのイメージ

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システム（以降，提案システムと呼ぶ）とは，複数大画面を使った協調作業において，ユーザがカメラ付き装着型携帯端末を用いて大画面とジェスチャインタラクションを行い，情報共有できるシステムである（図 2.1）．

以下に，提案システムの利点について説明する．

直感的な操作性が得られる

本システムでは大画面とインタラクションを行う際に両手が空いているため，両手によるのハンドジェスチャが可能になり，直感的な操作性が得られる．

また，ジェスチャ認識がユーザ側で行われるため，インタラクションを行う際に場所や大画面に拘束されず，自由に動き回ることができる．

情報共有ができる

ユーザがハンドジェスチャだけで大画面から情報を取得したり，また，携帯端末内の情報を大画面上で提示したりできる．そして，携帯端末によってユーザの識別が容易になり，複数人で同時に情報共有が出きる．

コンピュータを意識しない

ユーザが大画面と情報共有する際に，自分の携帯端末を手で持って，指で操作する必要がないため，コンピュータを意識することがなく，情報共有に集中できる．

2.3 想定する利用シナリオ

建設会社で働く A さんは，数十人のクループであるプロジェクトに取り組んでいた．締め切りが近いある日，自分が設計した図面に問題があることに気づき，短期間で再設計することになった．しかし，設計がうまくいかず，他の同僚たちと意見交換をすることにした．そこで，A さんが自分の作業大画面から壁面上の巨大な共有大画面に設計中の図面をハンドジェスチャによって移動させ，同僚たちから意見を聞いた．すると数人の同僚がそれぞれの作業場からハンドジェスチャだけで，図面を回転させたり，拡大縮小したりして議論し始めた．同僚たちがハンドジェスチャによって図面上で書き込んだり，別の図面を隣に表示して比較したり，図面を自分の作業大画面に移動させて確認したりして様々な議論が行われた．また，仕事を終えて変える途中だった同僚が「家で確認したい」と言い，大画面上の図面や同僚たちが書き込んだメモや図をハンドジェスチャによって自分の携帯情報端末内にコピーした．その後，同僚たちとの情報共有から多くの意見を得られた A さんが締め切りの前に図面を完成させることができた．

このシナリオでは，マウスやタッチ操作ができない壁面上の巨大な大画面と離れた場所からハンドジェスチャによってインタラクションできることが示されている．また，同時に複数人で大画面とインタラクションを行い，大画面上の情報に対して，回転，拡大縮小，移動，描画など様々な操作ができることが示されている．そして，複数の大画面の間や大画面と自分の情報端末の間で情報を自由に移動できることが示されている．また，情報共有の際にユーザが自分の携帯端末を操作しないため，コンピュータを意識せずに，ジェスチャだけで全ての操作が行えることが示されている．

第3章　プロトタイプシステム「MobaJes」

3.1　システムの概要

本研究では提案システムの実現に向けてプロトタイプシステムを実装した．そして，このシステムの 1 つのアプリケーションの例として，大画面を使って画像情報共有システムを作成した．このシステムでは，ユーザがカメラ付き Android 携帯端末をペンダントのように身に付け，ハンドジェスチャをすることで大画面とインタラクションでき，自由に情報共有できる．また，ハンドジェスチャにより大画面上で情報を操作するができる．

3.2　インタラクション

ここでは，本システムの具体的なインタラクションについて説明する．

3.2.1　基本ジェスチャ

ユーザが大画面とインタラクションを行う際に「指す」,「掴む」,「放す」といった次の 3 つの基本ジェスチャを使う．そして，これらのジェスチャを両手を用いて行うことで様々なインタラクションができる．

図 3.1: 指す

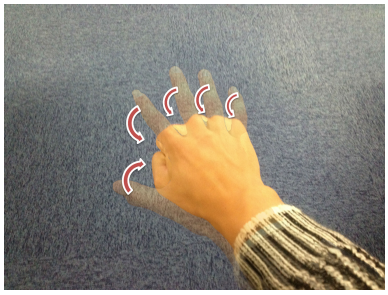

図 3.2: 掴む

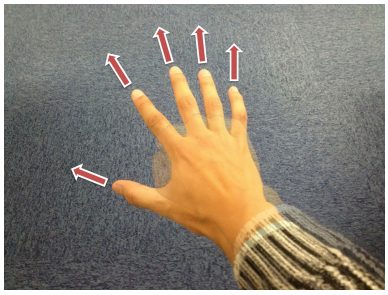

図 3.3: 放す

- 指す　：　ユーザが人差し指で指すようなジェスチャで大画面上で情報の選択することができる（図 3.1）.

- 掴む　：　ユーザが情報の上にカーソルを合わせて，手を握るジェスチャをすることで情報をドラッグすることができる（図 3.2）.
- 放す　：　ユーザが情報をドラッグしている際に手を開くジェスチャをすることで情報を放すことができる（図 3.3）.

3.2.2　大画面上でのインタラクション

ユーザは大画面とインタラクションしている際に，大画面上の情報に対して次のインタラクションが可能になっている.

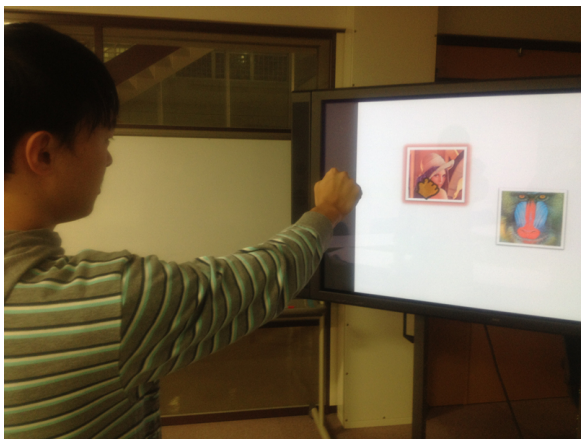

図 3.4: 配置の変更

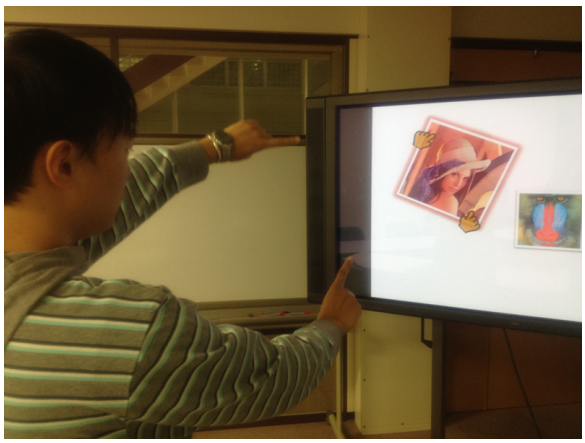

図 3.5: 表示サイズや傾きの変更

図 3.6: 取得

図 3.7: 提示

- 配置の変更　：　ユーザが情報に対して「掴む」ジェスチャをして，別の配置に移動させて「放す」ジェスチャをすることで情報の配置を変更することができる（図 3.4）.
- 表示サイズや傾きの変更　：　ユーザが情報に対して両手で「指す」ジェスチャをして，両手を離すことでその表示サイズを拡大し，両手を近づけることで縮小することができる．また，両手を回すことで情報を回転させることができる（図 3.5）.
- 取得　：　ユーザが大画面上の情報に対して 2 秒間以上「指す」ジェスチャをすることで情報を携帯端末に取得できる（図 3.6）.
- 提示　：　ユーザが片手の 3 本指のジェスチャをして画面の下の部分に携帯端末内の情報のサムネールを表示させ，その中の情報に対して 2 秒間以上「指す」ジェスチャをすることで，情報を大画面上で提示できる（図 3.7）.

第4章　システムの実装

このシステムではユーザがカメラ付き Android 携帯端末（以降，携帯端末と呼ぶ）を装着してインタラクションを行う．装着の方法は，携帯端末のカメラがユーザの前方を向く姿勢で，ペンダントのように胸の辺りに装着する．携帯端末を装着している様子を図 4.1 で示す．

図 4.1: 携帯端末を装着している様子

4.1　開発環境と言語

開発環境として Eclipse 3.7.1 を使用し，大画面側や携帯端末側のアプリケーションを Java 言語でプログラミングした．携帯端末として Android OS 2.3 を搭載した，Samsung Google Nexus S[1] という Android[2]端末を使用した．携帯端末側のアプリケーションは AndroidSDK[3]を用いて実装した．画像解析には Android 用 OpenCV[4] ライブラリを用いた．

[1]http://www.google.com/phone/detail/nexus-s
[2]http://www.android.com/
[3]http://developer.android.com/sdk/index.html
[4]http://opencv.jp//

4.2　システム構成

本システムは，Android 携帯端末や PC に接続された大型液晶ディスプレイ，無線通信が可能なネットワーク環境からなる（図 4.2）.

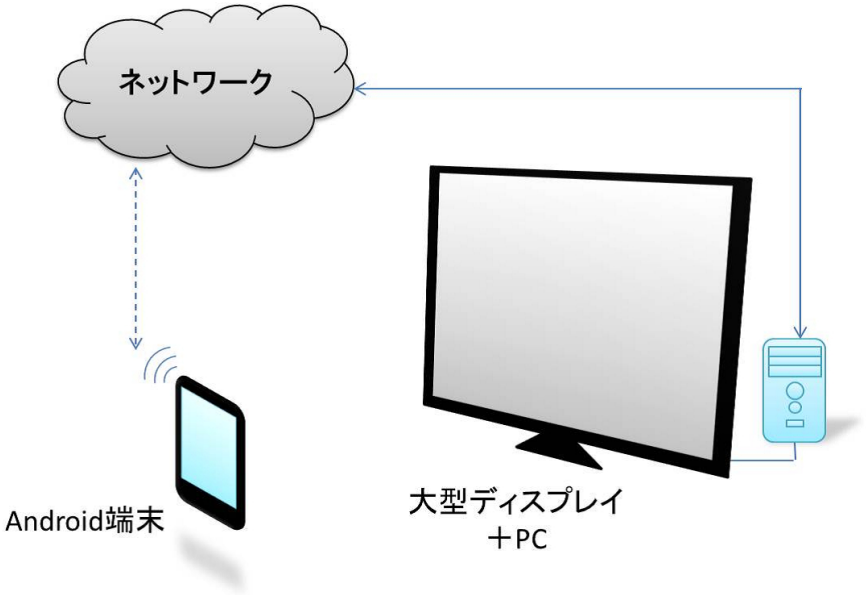

図 4.2: システム構成

4.3　通信方式

本研究では通信方式としてソケット通信を使用して携帯端末と大画面の間で通信を行う.

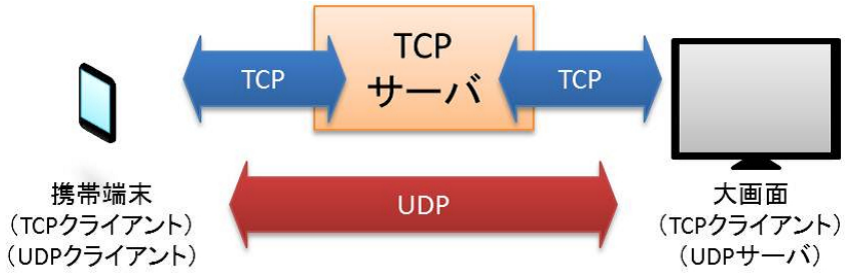

図 4.3: 通信方式

通信するメッセージの内容によって，TCP プロトコルと UDP プロトコルを使い分けて通信している（図 4.3）．

携帯端末からのジェスチャの座標情報がリアルタイムで送信される必要があるため，UDP プロトコルを用いて直接大画面に送信させている．

そして，コマンドや情報の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，携帯端末と大画面がクライアントとして 1 つのサーバに接続されており，サーバが複数のクライアントからのメッセージを分析して，適切な送信先にメッセージを送信する．

以下に，TCP プロトコルを用いて携帯端末と大画面の間で行われる通信の内容を説明する．

コマンドの送信 ： 携帯端末でジェスチャが確定したら，確定したコマンドの情報をサーバに送信する．するとサーバが送信先の大画面にその情報を送信する．大画面はコマンドを受け取ったら画面上のカーソルをコマンドに対応するカーソルに変更する．

情報の取得 ： 携帯端末に情報を取得させるとき，まず携帯端末から情報取得のコマンドとそのジェスチャの座標情報が送信される．大画面が情報取得のコマンドを受け取ったらそのコマンドの座標にある情報を携帯端末に送信する．

情報の提示 ： 情報提示を行うとき，まず携帯端末内の情報のサムネールを大画面に送信し，大画面上で表示させる．ユーザがその中から情報を選択すると，携帯端末内から選択されたサムネールの元の情報が大画面に送信される．

4.4　大画面側アプリケーション

大画面側では，画像情報を表示させ，様々な操作を行うことができるアプリケーションを実装した．

4.4.1　GUI の実装

アプリケーションの GUI は Java の Swing を用いて作成した．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．アプリケーションの GUI の例を図 4.4 で示す．

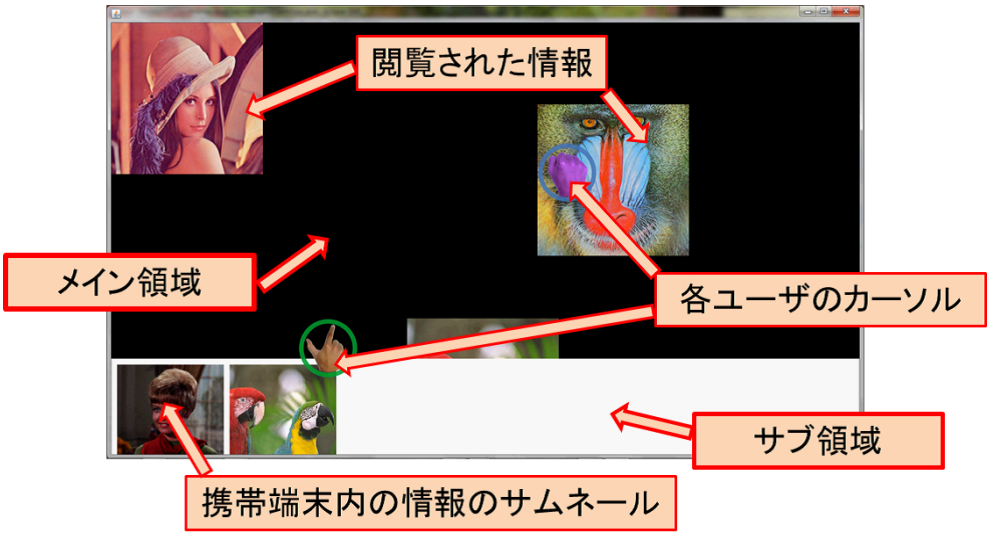

図 4.4: 大画面側アプリケーションの GUI

GUI は，まずメイン領域があり，その上に，情報やユーザのカーソルが表示される．また，携帯端末内の情報のサムネールを閲覧するサブ領域がメイン領域の下の部分に表示される．

4.5　携帯端末側アプリケーション

携帯端末ではユーザが行ったジェスチャを認識して，ジェスチャの位置情報，コマンドを送信するアプリケーションを実装した．

4.5.1　ジェスチャを管理するクラスの作成

本研究でハンドジェスチャを扱う Gesture クラスを作成した．また，ジェスチャの種類を管理する GestureType クラスと，両手を管理する Hand クラス，手の指を管理する Finger クラス，指先の点を管理する Tips クラスを作成した．クラス間の関連を図 4.5 で示す．

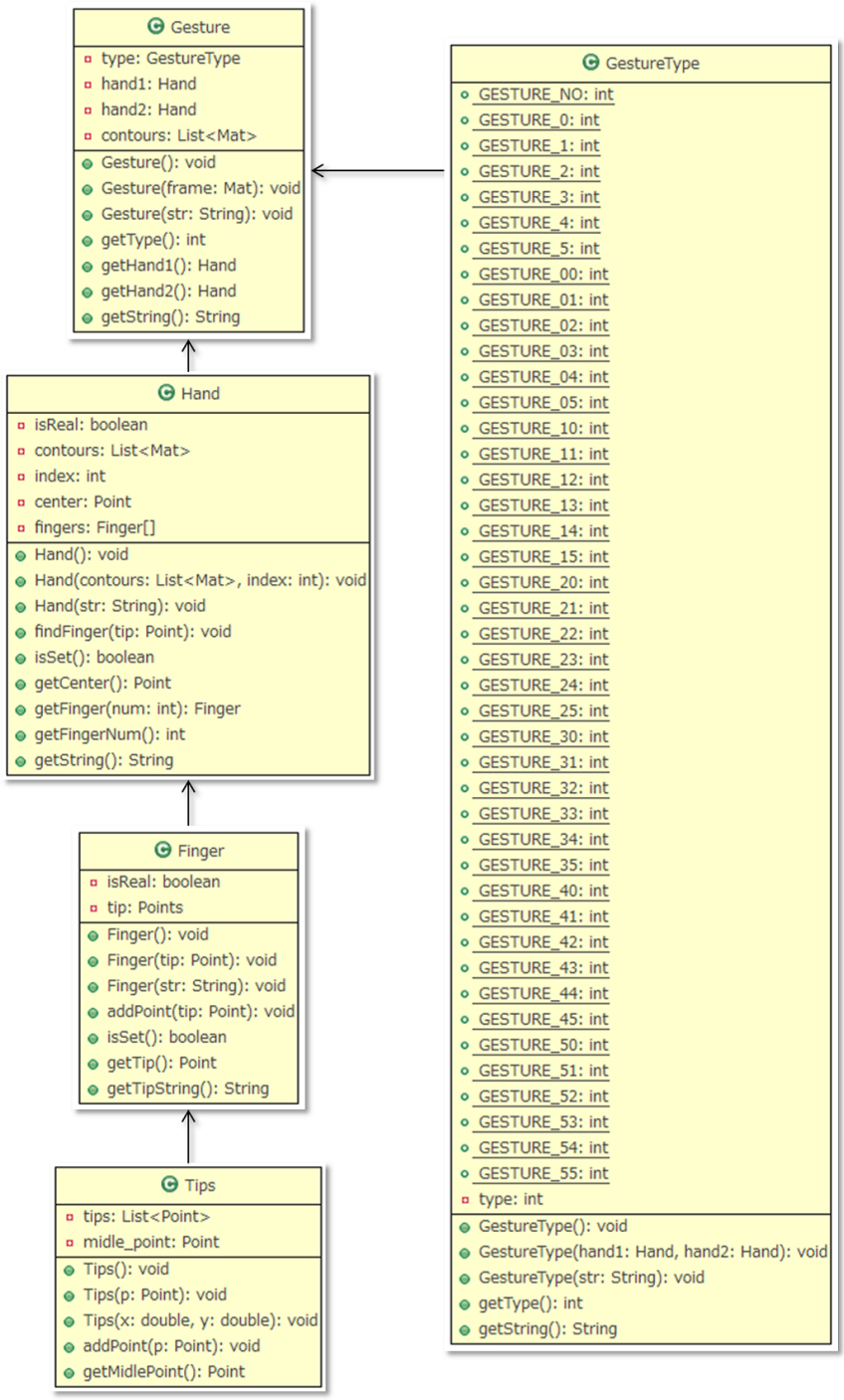

図 4.5: ジェスチャを管理するクラスの関連

4.5.2 ジェスチャ認識

ジェスチャ認識は次の 3 つの処理を順次に行うことで行っている．

1. 携帯端末のカメラフレーム frame から Gesture クラスのメソッド Gesture(Mat frame) を用いて肌色領域を検出する．
2. そして，見つかった領域の輪郭の情報から Hand クラスの Hand(List<Mat> contours, int index) メソッドを用いて指先を検出する．
3. 最後に，ジェスチャを判別する．

以下に，これらの処理をより詳しく説明する．

肌色領域の検出

最初に Android 端末のカメラから入力された画像（図）は YUV420 色空間で取得される．この画像の各画素が肌色であるかどうかをチェックするために，画素情報がより単純な並びをしている RGB 色空間か HSV 色空間に変換する．筆者は HSV 色空間で処理する．変換の際に OpenCV の色変換関数（4.1）（4.2）を用いて次のように変換する．まずは，YUV420 色空間から RGBA 色空間に変更する．次に RGB 色空間から HSV 色空間に変更する．関数の引数はそれぞれ入力画像の行列，出力画像の行列，どの色空間からどの色空間に変更するかを表す定数，出力画像のチャンネルを表す．

$$Imgproc.cvtColor(mYuv, mRgba, Imgproc.COLOR_YUV420sp2RGB, 4) \tag{4.1}$$

$$Imgproc.cvtColor(mRgba, mHsva, Imgproc.COLOR_RGB2HSV, 4) \tag{4.2}$$

そして，HSV 色空間に変更された画像の中から（4.3）の式を使い肌色の領域だけを抜けとったマスク画像（図）を作成する．この式は肌色検出に関する論文 [2][3][4][5] から参考にして，独自に予備実験を行った結果から得られたものである．

$$0 < H < 20 \qquad 38 < S < 192 \qquad 89 < V < 243 \tag{4.3}$$

その際に画素の最低値と最大値を指定してマスク画像を作成する関数（4.4）を用いる．関数の引数はそれぞれ入力画像の行列，画素の最低値，画素の最高値，出力されるマスク画像の行列である．この関数は入力画像の各画像をチェックして，画素の値が指定された範囲内にあれば画素を 1，そうでなければ画素を 0 とする同じ大きさの 2 値画像を出力する．

$$Core.inRange(mHsva, newScalar(0, 38, 89, 255), \\ newScalar(20, 192, 243, 255), mMask) \quad (4.4)$$

最後にノイズ除去のためにメディアンフィルタ（4.5）] を用いて平滑化する．関数の引数はそれぞれ入力画像の行列，出力画像の行列，メディアンフィルタの正方形の辺の長さである．

$$Imgproc.medianBlur(mMask, mMask, 9) \quad (4.5)$$

図 4.6: 肌色領域に検出.a) 元画像, b) 肌色領域の抽出, c) ノイズ除去後の画像.

手とその重心点の検出

肌色領域の検出で得られた 2 値画像から肌色領域の輪郭を抽出する．その際に輪郭を抽出する関数（4.6）を用いる．関数の引数はそれぞれ，入力 2 値画像の行列，輪郭を格納用行列のリスト，輪郭を抽出するときに階層情報を格納する行列のリスト，出力される輪郭の階層を指定する定数，抽出する輪郭の種類を指定する定数となっている．ここでは，階層情報を必要ないため Imgproc.RETR_LIST を指定している．

$$Imgproc.findContours(mMask, contours, hierarchy, Imgproc.RETR_LIST, \\ Imgproc.CHAIN_APPROX_SIMPLE) \quad (4.6)$$

そして，次のような条件で手の領域を抽出する．

- 領域の面積が 1200 以上である
- 領域の面積が最も大きな領域，あるいはその次の大きな領域

そして，見つかった領域を手の領域だと判断し，その数を手の数とする．そのとき輪郭から領域の面積を求める関数（4.7）を用いる．引数に 1 つの輪郭を示す行列を指定すれば，出力としてその面積が戻る．ここでは輪郭のリストから i 番目の輪郭を指定して，その面積を求めている．

$$Imgproc.contourArea(contours.get(i)) \tag{4.7}$$

次に，手の領域の重心点を次の関数（4.8）を用いて求める．引数に 1 つの輪郭を示す行列を指定すれば，出力としてその重心点が戻る．重心点の型は OpenCV の Point 型である．

$$Imgproc.moments(contours.get(i)) \tag{4.8}$$

図 4.7: 手とその重心点の検出.a) 輪郭の抽出, b) 面積によるフィルタリング, c) 領域の重心点.

指先の検出

最後に，抽出された 1 つ，あるいは 2 つの輪郭から指先の条件を満たす形状の部分指先だと判断する．そして，指先である条件は以下のように設定する．

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（4.8-a）．2 つ目の条件により指同士の間の谷が検出されなくなる．そして，本研究では k を 20 とした場合に指の認識率が一番高かった．本来，この値は携帯端末のカメラの解像度や手とカメラの距離によって変化する値てある．今回の実装では定数として使った．

2 つ目の条件のため，指定された輪郭 contour の凸閉包を求める次の関数（4.9）を用いる．関数の引数はそれぞれ，指定された輪郭の行列と，出力される凸閉包が格納される行列である．

$$Imgproc.convexHull(contour, hull) \tag{4.9}$$

また，こ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 クラスを用いて 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（4.8-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（4.8-c）．

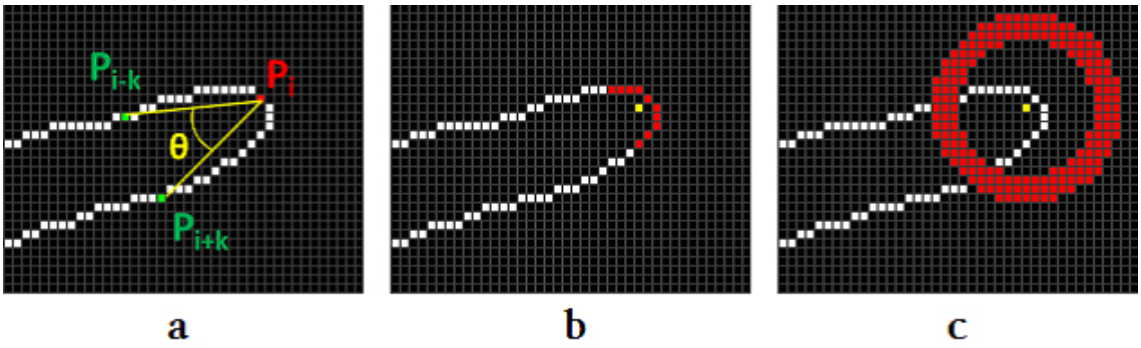

図 4.8: 指先検出.a) 凸型部分の検出, b) 指先の候補点の集合, c) 指先判定.

ジェスチャ判別

手やその指を求めた後，検出された手の数や指の数のコンビネーションでジェスチャを判断する．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 まであるため 6 種類のジェスチャを認識できる．両手のジェスチャでは，指の数がそれぞれ 6 本とし，合わせて 36 通りのジェスチャを認識できる．

4.5.3 大画面との接続

大画面との接続にはユーザの向いている方向情報を用いる．そのため，最初にユーザが向いている方向情報をサーバに送信する．そして，サーバがその情報を登録されている大画面の方向情報と比較して大画面を選択する．

ユーザが向いている方向の検出

ユーザが向いている方向は Android 端末に搭載されている磁力センサの値から取得することができる．

そのために Android 端末上のアプリケーションに SensorEventListener インターフェースを実装し，onSensorChanged(SensorEvent event) メソッドを用いて各センサの値を取得する．そして取得される event.values[0] が Android 端末の向いている方向情報である．この値は携帯端末の Y 軸が地球の磁北から時計回りに見たなす角度であり，本研究のように Android 端末の X 軸が上を向くような姿勢では実際の方向より 90 度低い値が得られる．そのため，サーバに送信する際はこの値に 90 を加算し (式 4.10)，結果が 0 から 360 の間にあるように調整する (式 4.11).

$$\alpha = \alpha + 90; \tag{4.10}$$

$$\alpha = (\alpha < 360)?\ \alpha : \alpha - 360; \tag{4.11}$$

携帯端末を画面が上を向くように置いた状態での各軸の方向を図 4.9 で示す.

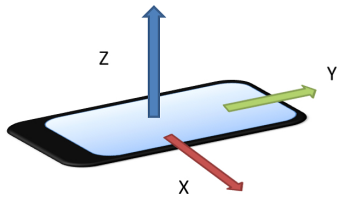

図 4.9: 携帯端末の X,Y,Z 軸

大画面の選択

大画面の選択は磁北を基準にユーザの向いている方向と逆方向（180 度回転させた方向）を大画面の向く方向として，サーバ内で登録されている大画面の方向情報と照らし合わせることで行われる．そして誤差は± 30 度とする．例えば図 4.10 では，ユーザの方向情報 α からその逆方向を求め，それと一番近いのが θ_3 であるため，大画面 3 が選択されることが示されている．

もし登録されている大画面の中に指定された範囲内の方向値を持つ大画面が存在しない場合はユーザが大画面に向いていないと判断し，画面選択を行わない．また，登録されている大画面の中に指定された範囲内の方向値を持つ大画面が複数存在する場合は指定の範囲の中央値に最も近い画面が選択される．

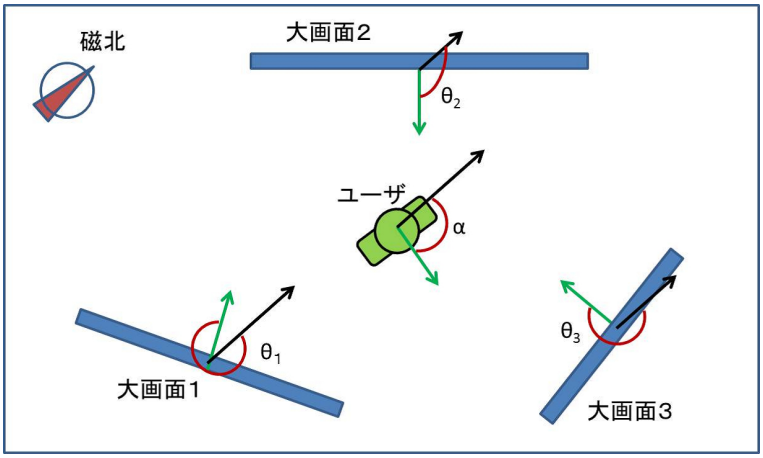

図 4.10: 大画面の選択（部屋の真上から見た場合）

第5章　関連研究

5.1　携帯端末を用いた大画面インタラクション

Ballagas ら [6] は，携帯端末のカメラを使い，大画面上のオブジェクトを選択し，インタラクションを行うシステムを提案した．このシステムは，大画面上で複数のマーカを表示させ携帯端末に読ませることで携帯端末で映している大画面の座標を検知し，オブジェクトを特定することで選択できる．またオプティカルフローアルゴリズムを用いて携帯端末の動きに大画面上のカーソルを連動させることで大画面とインタラクションできる．また，Zhong ら [7] は，カメラ付き携帯端末のカメラ情報からその 3 次元動きを追跡する方法を提案し，複数人で画面上で 3 次元曲線を描くことができるシステムを開発した．Jeon ら [8] は，カメラ付き携帯端末を用いてモーションフローやマーカトラッキングを行い，大画面上のオブジェクトと操作できるシステムを提案した．Boring ら [9] は，携帯端末のカメラで，実世界のディスプレイを映すことで，カメラ映像越しにディスプレイ上の情報を操作し，また複数ディスプレイ間で情報を移動させることができるシステムを開発した．

これらの研究は，携帯端末を用いて大画面とインタラクションできるところが本研究と類似しているが，ユーザがインタラクションを行う際に携帯端末を手で持つ必要があり，長時間利用する場合ユーザの手に負担を掛けてしまう問題がある．また，携帯端末と大画面の間で情報共有を行いっていない．本研究では，携帯端末を手で持たずに体に装着し，ハンドジェスチャによって大画面とインタラクションを行うため，ユーザにかかる負担が少なく，大画面との情報共有も可能になっている．

5.2　ハンドジェスチャを用いた大画面インタラクション

Argyros ら [10] 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．Clark ら [11] は，赤外線 LED を手に取り付け，それを赤外線カメラで読み取ることでハンドジェスチャを認識し，画面とインタラクション行うシステムや，携帯端末のタッチスクリーンをタッチパッドのように使うことで画面上のカーソルを移動させ，音声認識ツールを用いてコマンドを送ることができるシステム，デプスセンサを用いてハンドジェスチャを認識し画面とインタラクションできるシステムなど手が届かない場所にある画面に対してインタラクションするいくつかの手法を提案した．Shi ら [12] 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．

これらの研究はカメラやデプスセンサを用いてハンドジェスチャを認識し，大画面上のカーソルをコントロールするところが本研究と類似している．しかし，ジェスチャを認識するためのカメラが固定されており，インタラクションを行う際に自由に移動できない．本研究は装着型携帯端末でジェスチャを認識しているため，インタラクションの際に場所に拘束されず自由に動くことができる．

5.3 大画面との情報共有

Jin ら [13] は，公共場において携帯端末のカメラで壁面上の QR コードを読み込むことで，携帯デバイスを壁面に接続し，壁面と携帯端末の間で情報共有できるシステムを提案した．この研究では，携帯端末のディスプレイを操作面とすることで大画面上のカーソルをコントロールすることができ，独自に開発した情報提示方法を用いて大画面と情報共有を行っている．また，Bragdon ら [14] は，デプスセンサや携帯端末，タッチスクリーンを用いて専用の空間内で少人数で効率的な協調作業を行えるシステムを提案した．

これらの研究は携帯端末を用いて大画面とインタラクションしながら，大画面と情報共有ができる点で本研究と類似している．しかし，これらの研究が大画面と情報共有を行う際に携帯端末を手で持つ必要があるため，長時間使う場合ユーザに負担を掛けてしまう問題がある．本研究は，ユーザへの負担が少ないハンドジェスチャを用いて大画面と情報共有できるところがこの研究と異なっている．

5.4 装着可能なジェスチャインタフェース

Gandy ら [15] は赤外線 LED や赤外線カメラを用いてハンドジェスチャを認識する小型デバイスを開発した．ユーザがデバイスをペンダントのように胸の辺りにつけて，その前に両手でハンドジェスチャすることで周辺の電子機器を操作するシステムを提案した．装着型デバイスによりジェスチャを認識できるところが本研究と関連するが，デバイスの機能ががカメラ画像を取得して，送信することに特化しており，情報共有ができない．本研究が，ジェスチャ認識デバイスとして多機能な携帯端末を用いることで情報共有を可能にしているところが異なる．

Kratz ら [16] は，赤外線距離センサを搭載した小型デバイスを開発し，それを携帯端末に取り付けて，手を近くに動かすことで携帯端末をコントロールできるシステムを開発した．この研究は，携帯端末に直接触れず，ハンドジェスチャするだけで携帯端末を操作できるところが本研究と類似している．しかし，赤外線距離センサで計測できる距離が短いため，ジェスチャを携帯端末の近くに行う必要がある．本研究は，携帯端末のカメラを使っているためより広い空間でジェスチャを行うことができる．

Mistry ら [17] は装着可能なコンピュータとカメラや小型プロジェクタを組み合わせたデバイスを用いて実世界に情報を重畳表示しながら，ハンドジェスチャによりインタラクションを

行うことができるシステムを提案した．この研究は装着型ハンドジェスチャインタラクションシステムであることや情報共有を行うことができるというところが本研究と類似するが，インタラクションの際に指にマーカを付ける必要がある．本研究は手に何も付けずにジェスチャを認識できるところがこの研究と異なる．

Harrison ら [18] は，近距離デプスセンサと小型プロジェクタ，装着可能なコンピュータ用いて壁や本，手の平などあらゆる平面をタッチ操作可能な平面にするシステムを開発した．携帯型ジェスチャインタラクションが可能であり，手にマーカなどを付けずにハンドジェスチャを認識しているところが本研究と類似している．しかしこのシステムでは，インタラクション行うためにユーザが手で平面に直接触れる必要である．本研究ではユーザが平面を必要とせず，空中でジェスチャを行うことができるところがこの研究と異なる．

第6章　議論

6.1　試用から得られたシステムの利点と問題点

筆者が本研究の提案システムのプロトタイプシステム「MobaJes」実装し，試用した．そして，その結果として得られた利点や問題点について以下に述べる．

本システムを使用することで，手にマーカやセンサを付ける必要がなく，両手によるハンドジェスチャを行うだけで大画面とインタラクションすることができた．また，大画面から離れた場所からでもインタラクションできることを確認した．手で指す，握る，放すなど自然なジェスチャを行うことで情報にたいして選択，移動，回転，拡大縮小などの操作を行うことができた．したがって，本システムではハンドジェスチャインタラクションにより，大画面インタラクションとして自由度の高い操作性が得られたと考えられる．

また，ハンドジェスチャだけで大画面から情報を取得し，また大画面へ情報を提示することができ，大画面と情報共有できることがわかった．特にこのシステムが，複数人で複数大画面を使って情報共有を行う際に有効だと感じた．

しかし，インタラクション中にハンドジェスチャが認識されなくなる問題があった．そして，この問題の原因は２つあった．その１つは，装着型携帯端末のカメラでジェスチャを認識できる範囲が小さかった．この問題の解決策として，携帯端末のカメラにより高角度のレンズを取り付け，カメラが取得できる範囲を広げることができると考えられる．また，装着型携帯端末を胸の辺りではなく頭に装着することでより広範囲でジェスチャを認識できると考えられる．しかしその場合，ユーザに負担がかかってしまう恐れが考えられるため，装着型携帯端末の装着位置について再検討する必要がある．

もう１つの原因としては，ジェスチャの認識精度に問題があった．本研究では，肌色情報を用いてジェスチャを認識しているため，使用する環境の光の強さの変化で，肌色も変化してしまい，ジェスチャを認識できなくなることがあった．この問題を解決するために，今回の実装で定数として扱っていたHSV色空間置ける肌色の閾値を環境の光にあわせて動的に変動させるアルゴリズムを実装することが考えられる．また，デプスカメラを用いて環境の光の影響を受けにくい，よりロバストな方法でジェスチャを認識する方法が考えられる．

そのほかに，各大画面が方向情報だけで識別されているため，大画面が同じ方向で並んでいるような環境では大画面の選択ができないという問題があった．この問題を，ARToolKitを用いて，大画面上で表示されたマーカによって大画面を識別させることで，解決できると考えられる．

6.2　今後の展望

今後の展望としては，ハンドジェスチャによって大画面上で描画できる機能を実装することが挙げられる．また，大画面上で扱える情報の種類を増やすことが考えられる．例えば，映像やテキストファイル，ウェブページなど．そして，それらを操作するためにジェスチャの種類も増やす予定である．

また，音声認識の技術を用いて，大画面上で文字入力できるようにすること．

第7章　結論

本研究では複数大画面を使った協調作業において，大画面と直感的にインタラクションしながら，コンピュータを意識せずに情報共有できる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「MobaJes」システムを開発した．

本プロトタイプシステムでは，ユーザがカメラ付き装着型携帯端末を装着し，その前で両手によるハンドジェスチャを行うことで，大画面と直感的にインタラクションができる．そして，ハンドジェスチャだけで大画面上の情報に対して選択やドラッグドロップ，拡大縮小，回転，位置変更などの操作を行い，大画面とユーザの携帯端末の間で情報を移動させることが可能になっている．また，ユーザが自分の携帯端末を手で操作する必要がないため，コンピュータを意識せずに情報共有を行うことができる．

今後は，被験者実験によってシステムを客観的に評価し，そのフィードバックからシステムの改善を行っていきたい．

謝辞

本論文を執筆するにあたり，指導教員の田中二郎先生をはじめ，志築文太郎先生，三末和男先生，高橋伸先生にはゼミやミーティングを通して大変貴重なご意見やご指摘を頂きました．心から感謝致します．また，田中研究室の皆様にはゼミや学校生活の中で数々のご意見をいただきました．心より感謝いたします．そして最後に，私に日本で勉強できる機会を与え，経済面で支えてくれた日本の政府や，大学生活を送る中，精神面で支えてくれた両親や友人，そして，日本での生活の様々な面でお世話になった全ての人に心より感謝いたします．本当にありがとうございました．

参考文献

[1] 教育機器編集委員会. 産業教育機器システム便覧. 日科技連出版社, 1972.

[2] 松橋聡, 藤本研司, 中村納, 南敏. 顔領域抽出に有効な修正 HSV 表色系の提案. テレビジョン学会誌, 社団法人映像情報メディア学会, Vol. 49, No. 6, pp. 787-797, 1995.

[3] 呉海元, 陳謙, 谷内田正彦. ファジイパターン照合を用いた色彩画像からの顔検出システム. 電子情報通信学会論文誌. D-II, 情報・システム, II-情報処理, 社団法人電子情報通信学会, Vol. 80, No. 7, pp. 1774-1785, 1997.

[4] 本郷仁志, 山本和彦. 動領域内の肌色推定による顔領域および顔部品抽出. 映像情報メディア学会誌 : 映像情報メディア, 社団法人映像情報メディア学会, Vol. 52, No. 12, pp. 1840-1847, 1998.

[5] Mohamed-Ikbel Boulabiar, Thomas Burger, Franck Poirier, and Gilles Coppin. A low-cost natural user interaction based on a camera hand-gestures recognizer. *Proceedings of the 14th international conference on Human-computer interaction: interaction techniques and environments - Volume Part II*, pp. 214-221, 2011.

[6] Rafael Ballagas, Michael Rohs, and Jennifer G. Sheridan. Sweep and point and shoot: phonecam-based interactions for large public displays. *CHI '05 extended abstracts on Human factors in computing systems*, pp. 1200-1203, 2005.

[7] Yu Zhong, Xin Li, Mingming Fan, and Yuanchun Shi. Doodle space: painting on a public display by cam-phone. *Proceedings of the 2009 workshop on Ambient media computing*, pp. 13-20, 2009.

[8] Seokhee Jeon, Jane Hwang, Gerard J. Kim, and Mark Billinghurst. Interaction with large ubiquitous displays using camera-equipped mobile phones. *Personal Ubiquitous Comput.*, Vol. 14, No. 2, pp. 83-94, February 2010.

[9] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: mobile interaction through video. *Proceedings of the 28th international conference on Human factors in computing systems*, pp. 2287-2296, 2010.

[10] Antonis A. Argyros andManolis 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.

[11] Adrian Clark, Andreas Dnser, Mark Billinghurst, Thammathip Piumsomboon, and David Altimira. Seamless interaction in space. *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, pp. 88-97. 2011.

[12] Junhao Shi, Mingmin Zhang, and Zhigeng 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.

[13] 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.

[14] 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.

[15] Maribeth Gandy, Thad Starner, Jake Auxier, and Daniel Ashbrook. The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. *Proceedings of the 4th IEEE International Symposium on Wearable Computers*, pp. 87-94, 2000.

[16] Sven Kratz and Michael Rohs. Hoverflow: exploring around-device interaction with IR distance sensors. *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1-4, 2009.

[17] Pranav Mistry, Pattie Maes, and Liyan Chang. WUW - wear Ur world: a wearable gestural interface. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pp. 4111—4116, 2009.

[18] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. OmniTouch: wearable multitouch interaction everywhere. *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 441-450, 2011.