

# 映像合成と画像認識を用いたマシンインターフェースの開発

システム科学技術学部 機械知能システム学科

2年 折戸 裕輔

指導教員 学部名 学科名

准教授 佐藤 和人

助教授 間所 洋和

## 1. 研究背景

前年度の自主研究において次世代マシンインタフェースの開発として手を動かすだけで画面を操作できるメガネを開発し、3次元CGモデルを現実存在するようにユーザが認識することを可能とするプログラムの研究を行った。

本年度の研究では前年度の自主研究を発展させ、複数のユーザが3次元CGモデルを認識でき、さらにプログラムを使用する複数のユーザが一つの3次元CGモデルを操作できるようにしたいと思い、本研究を行った。

## 2. 研究目的

前年度行った自主研究において制作したプログラムに以下の改良を加える。

- 手の認識をプログラムを変更し、グー・チョキ・パーのハンドサイン3種類とハンドサインを示した後の手の移動量も操作入力として利用できるようにする。
- 3次元CGモデルに対する描画位置の情報を、プログラムを使用する複数のユーザで共有し、同一の3次元CGモデルを見ているように改良する。

以上の2点をプログラムすることにより、マウスとキーボードを利用しないで3次元CGモデルを操作でき、複数のユーザから3次元CGモデルへ操作を受け付けるタンジブルユーザインタフェース(一つのインタフェースで複数の人が同時に作業できる。形のない情報を直接触れることができるようにした、より実体感のあるインタフェース)の開発を目指す。

## 3. 前年度自主研究の内容

次世代マシンインタフェースとして、ユーザが装着するメガネ型の液晶ディスプレイ上へ現実の映像に拡張現実(AR)プログラムである Parallel Tracking and Mapping for Small AR Workspaces(PTAM) © Isis Innovation Limited 2008 のプログラムを改良して、3次元CGモデルをリアルタイムで合成して表示し、その3次元CGモデルに、ユーザがグー・チョキ・パーのハンドサインを示すことで、3次元CGモデルの回転操作を出来るインタフェースの開発を行った。

## 4. 研究内容

### 3.1 ソフトウェアの開発

#### 1)開発環境及び実行環境

ソフトウェアの開発は、VisualStudio 2010 Professional を使い、C++言語で開発を行った  
実行環境として、前年度自主研究で作成したカメラ付きヘッドマウントディスプレイを2台制作した。(図1)



## 2) プログラミング

前年度自主研究で作成したプロジェクトを書き換える形で以下のプログラムを行った。

- i) 手の認識プログラムの改良
- ii) ネットワーク通信プログラム

### i) 手の認識プログラムの改良

ユーザの操作の幅を広げるために、ユーザが手でハンドサインを示した後に、ハンドサインを移動させた場合に画像処理ライブラリ **OpenCV** とラベリング処理ライブラリ **cvblob** を組み合わせたプログラムによってハンドサインの形状が持つ重心の移動量からハンドサインの移動量を取得可能にし、ハンドサインを動かさず動作も操作の入力として使用可能にした。(図 2)

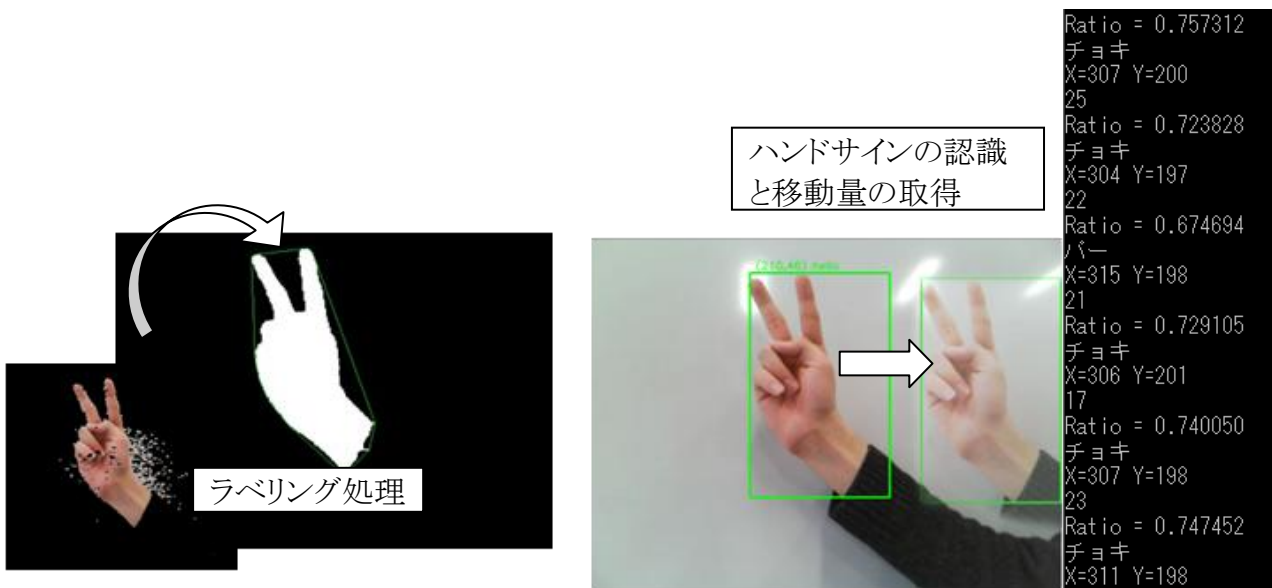


図 2 手のラベリング処理とハンドサインの認識

### ii) ネットワーク通信プログラム

C++言語においてネットワーク通信をサポートする API(Application Program Interface)である **Winsock2** を用いて、3次元 CG モデルの表示位置及びユーザが行った操作に関する情報をクライアントサーバモデルのネットワーク通信によって共有するプログラムの開発を行った。これにより複数のユーザが同一の 3次元 CG モデルを見ているかのように認識することが可能となった。(図 3)

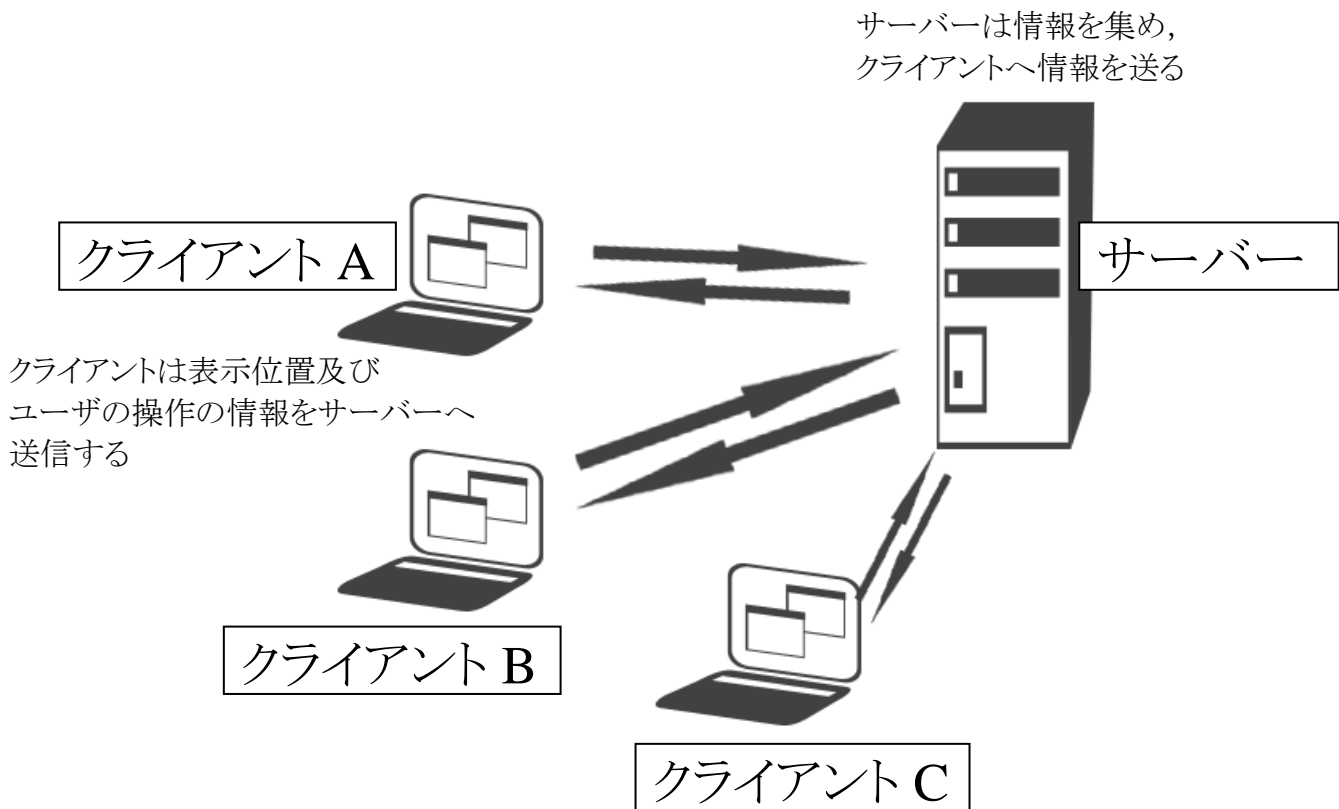


図 3 サーバークライアントモデル

## 5. 実験結果

ゲーのハンドサイン認識後のカメラが捕らえる画像上の上下, 左右の手の移動量を 3 次元 CG モデルの 3 次元空間上において指定した任意の平面 XY 座標の移動量に反映させることができた. (図 4)

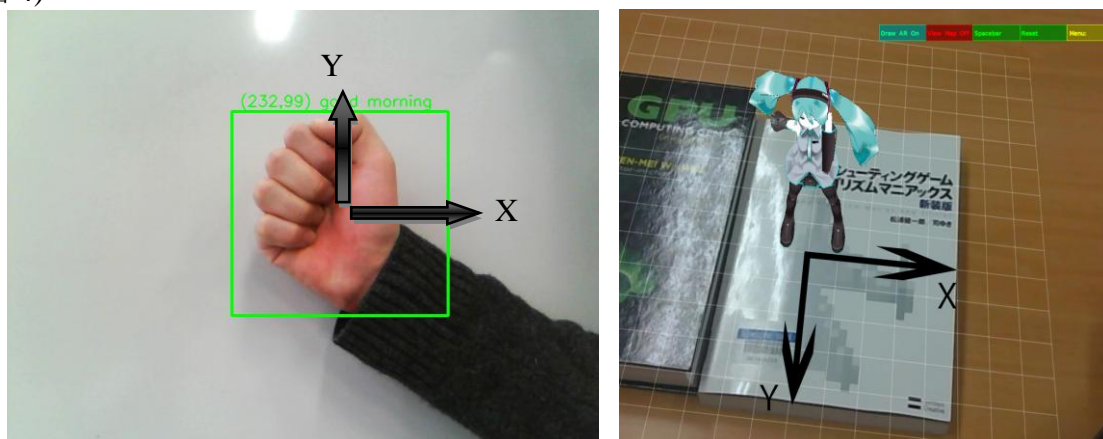


図 4 手の移動と対応関係

ネットワーク通信を用いたプログラム間のデータ転送に成功し, 3 次元 CG モデルの移動, 回転情報を共有することが可能となった. (図 5)

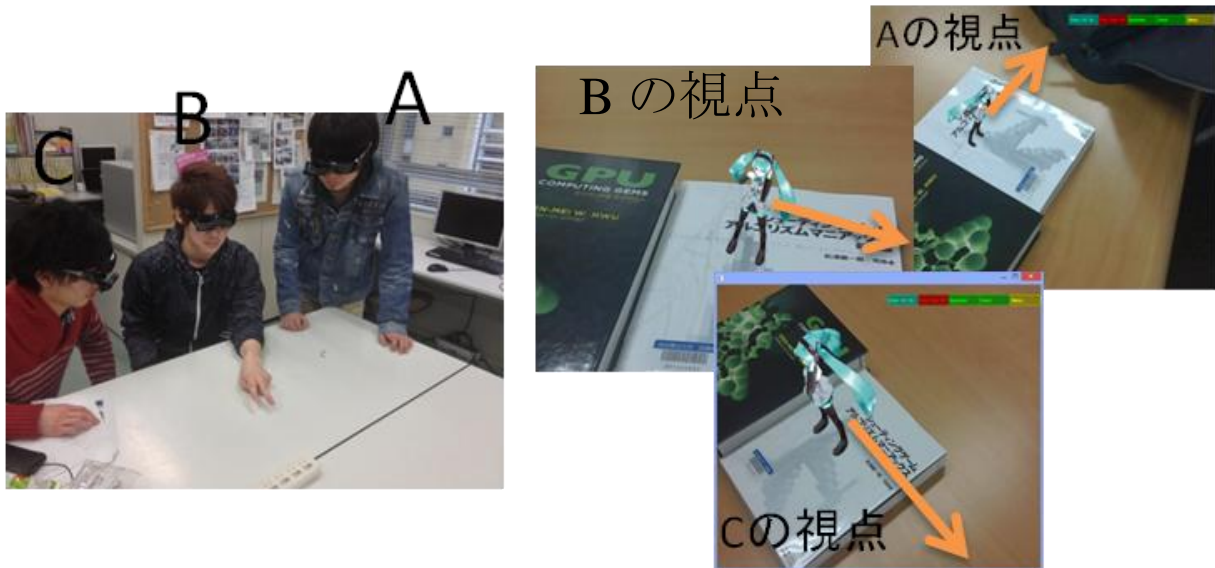


図 5 複数ユーザの利用の例

プログラム起動から 30 秒間のフレームレート(1秒間に 3 次元 CG モデルの描画を更新した回数)を計測したところ, 平均 6FPS(1 秒間に 6 回の更新)であった. 人の目においてある程度の滑らかな動きを認識させるためには 15FPS 以上必要であるが, 半分以下のフレームレートとなった

## 6. 研究成果

前年度プログラムが複数のユーザへ対応することが可能となった. ハンドサインの移動量もユーザの操作として扱えるようになった.

しかし, ユーザが行った操作の情報をネットワーク通信を用いて同期する処理によってコマ落ちを起こす現象が発生し, 映像をリアルタイムで表示できない問題が発生した. ユーザが快適に使用するためには自分の操作を可能な限り遅延なく反映させる必要があるため, この問題を今後の課題としたいと思った.

## 参考文献

[1]PTAM ISMR2007 掲載

[2]PTAM <http://www.robots.ox.ac.uk/~gk/PTAM/>

[3]OpenCV プログラミングブック 奈良先端科学技術大学院大学 OpenCV プログラミングブック制作チーム 中川信行 2007 年 10 月第二版, pp.196-217.

[4] cvblob <http://code.google.com/p/cvblob/>