

手を動かすだけで画面を操作できるメガネの開発

システム科学技術学部 機械知能システム学科

1年 折戸 裕輔

指導教員 システム科学技術学部 機械知能システム学科

准教授 佐藤 和人

助教 間所 洋和

1. 研究背景

SF映画やアニメなどに出てくる空中投影ディスプレイや、ユーザーがマシンに直接手を触れずにコンピュータを操作するシーンに感銘を受け、その仕組みを現在の技術の組み合わせにより擬似的に再現したいと思った。

据え置き型のカメラによるハンドサインの認識と、カメラの映像に3DCGが合成できるサンプルプログラムがあり、これらを組み合わせることで、再現できるのではないかと思った。

2. 研究目的

- ・画像中の手の認識
- ・カメラの映像より空間の認識

3. 研究内容

3. 1 ソフトウェアの開発

1) 開発環境

ソフトウェアの開発は、Visual Studio 2010 Professional を用いて、C++言語で開発を行った。また、ARを実現するためのプログラムは、現在複数のライブラリが開発されており、自主研究ではGeorg Klein氏とDavid Murray氏のPTAM (Parallel Tracking and Mapping for Small AR Workspaces) というプログラムを利用した。手の認識はOpenCV2.2.0を利用した。

2) 作成するソフトウェアの内容

本研究で開発したソフトウェアは、コンピュータに接続したUSBカメラで手の認識を行うと同時に、PTAMを用いた3DCGの表示を行うプログラムである。

3) PTAM

“小さなAR (Augmented Reality) 領域のための並列追跡とマッピング”

オックスフォード大学の完全子会社である「Isis Innovation」社のカメラから得た画像を

用いた SLAM (Simultaneous Localization and Mapping) によるマーカーレス AR 技術.

公式ホームページで配布されているサンプルプロジェクトを用いてライブラリを用意し、コンパイルできれば、マーカーレス AR を用意に再現できる。プログラム言語は C++ を使用し、CG の表示には OpenGL が用いられている。実行する場合、下位のライブラリに libcvd, gvars3, pthread, libjpeg, fltk, TooN, GLEW を用いる。

4) プログラミング

主に、PTAM のプロジェクトをベースに書き換えを行う形で開発を行った。

・ PTAM の最適化

PTAM のプログラムは、標準で利用できるカメラのリソースが IEEE1394 で接続されたカメラでなければ映像取得ができなかったため、EWCLIB を用いて UVC (USB Video Device Class) 対応の USB カメラから映像を取得できるようにした。次に、3DCG モデルをファイルからロードできるように GLMetaseq を追加した。この時、ファイルのロードが容易に行えるように Windows API を利用したファイルローダのプログラムを追加した。また、USB カメラを用いた PTAM の動作を図 1 に示す。

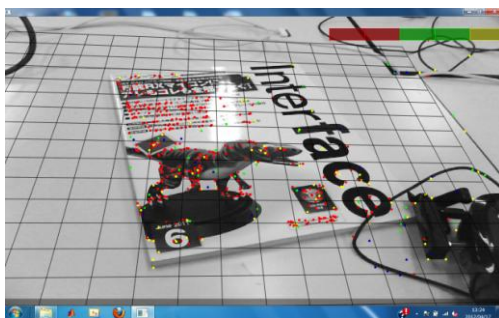


図 1. PTAM による空間認識

・ 手の認識

OpenCV のサンプルコードをベースにプログラムを作成した。ただし、このサンプルコードをコンパイルしたものは、研究室にある Windows Vista 以降の WindowsNT 系の OS の入ったコンピュータでは動作しなかった。原因は、OpenCV のカメラキャプチャが利用する Video for Windows-API の誤動作によってカメラキャプチャが正しく行われず、ポイント内のデータが空となるエラーがあったためだと思われる。そのため、EWCLIB で画像を取得し、OpenCV に渡すという処理で行うよう書き換えた。

・ プログラムの統合

「カメラから取得した画像上に 3DCG を合成するプログラム」、「手の認識プログラム」が揃ったので、統合を行った。このとき、カメラリソースの競合問題が発生した。PTAM では EWCLIB が、手の認識では OpenCV がそれぞれカメラドライバに対して画像取得を行おうとするため、双方のライブラリが互いに画像の取得を阻害していた。それを回避するため、カメラの画像取得をどちらか一方だけで行うようにした。これで、OpenCV と

PTAM の両方でカメラデータの共有ができるようになった。ここで、OpenCV による手の認識プログラムの結果を、PTAM の 3DCG 描画に反映させようと組み込んだ。

PTAM は、膨大かつ複数のソースコードによるプログラムで、いくつものクラス定義や複数のヘッダが各所で呼び出されており、PTAM 内に OpenCV のプログラムを組み込むと、二重インクルードや二重定義のエラーが発生した。通常ならば、ヘッダとソースコード全体の最適化を行うが、PTAM の改変は、自分の力量ではできないため、エラーが発生しないように各ソースコードのファイル内の隙間に組み込むことにした。OpenCV と PTAM において、共通して利用する EWCLIB のヘッダにより、二重定義のエラーが起きるので、カメラキャプチャを行なっているソースコードと同じファイル内の手の認証プログラムのソースコードを書き入れた。そして、手の認証プログラムのエントリーポイントを繰り返し呼び出して、動作させるように書き換えた。

以上により、手の認識と PTAM が同時に動作するようになった。

- ・ハンドサインとモデル動作の関連付け

手の認識の結果を 3DCG の動きに反映させるため、手の認識処理プログラムから 3DCG モデル表示処理へと手の認識結果を送るようにした。手の認証プログラムにより受け取ったデータを元に、モデルにアクションを起こさせた。現状ではハンドサインの認識しかできないため、グーならモデルを右に回転、チョキならモデルを左に回転させるようにした。プログラムの動作結果を図 2、3 に示す。



図 2. 開発したプログラムによる AR

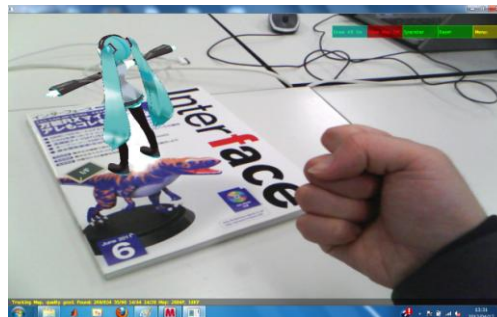


図 3. 開発したプログラムによる
モデルの回転操作

3. 2 ハードウェアの製作

作成したソフトウェアは、人の視界に情報を表示させることを目的とするため、人の視界を乳臭することと、視界に情報を重ねることを必要とする。

そのため、全面に USB カメラを固定した HMD (Head Mounted Display) を作成した。作成した HMD を図 4 に示す。

4. 研究結果

このプログラムは「カメラ画像からカメラの動きを計算するプログラム」、「CG モデルの描画プログラム」、「手の形状認識プログラム」など、行列計算、画像処理、画像合成、な

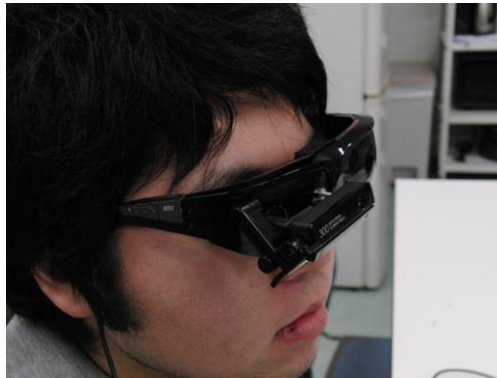


図4. カメラ付き HMD

どを同時にリアルタイムで行うため、実行するコンピュータに大きな負荷がかかる。そのため、研究室のコンピュータで動作させたところ、フレームレートが低くなり、手の認識においても認識精度が低く、誤認識が多かった。また、3DCGモデルを描画しているエリアに手をかざすと、手によりカメラの座標計算に誤差が生じてしまい、モデル描画が安定して行われないなどの問題が発生した。

5. まとめ

本研究では、カメラの映像にリアルタイムで3DCGを合成し、その3DCGに対しハンドサインを用いた移動、回転などの操作ができるソフトウェアを開発することで、全ての目的を達成することができた。しかし、以下の3つが今後の課題として残った。

・ハードウェアの性能の問題

作成したソフトウェアは、空間認識や3DCGの表示をリアルタイムに行うため、処理速度の速いCPUと、多くのメモリが必要となった。

・現実の情報の読み取り精度の問題

空間認識中にカメラを大きく動かすと、カメラが捉える瞬間の画像が大きくぶれてしまい、認識できなくなった。

・各認識プログラムの認識対象の競合の問題

空間認識している場所に手が被り、認識している空間が見えなくなると、正しく空間が認識されなかった。

参考文献

1. 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
2. PTAM ISMR2007 掲載
3. <http://kougaku-navi.net> 工学ナビ
4. http://www.geocities.jp/in_subaru/ewclib/index.html EWCLIB