

指動作認識を利用した スマートグラス上のユーザインターフェイス操作

森田 健太郎¹ 長田 剛典¹ 佐藤 健哉²

概要：近年、拡張現実 (Augmented Reality : AR) の技術が世の中に広まってきている。AR を表示するデバイスの一つにスマートグラスが挙げられるが、スマートグラス上に表示させた AR 情報を操作する手法に関しては、未だ確立されたものが無いのが現状である。そこで本研究では、赤外線を利用して指動作を認識することにより、スマートグラス上の情報を操作するシステムを提案する。

Operating User Interface on Smart Glasses by Recognizing Finger Movement

KENTARO MORITA¹ TAKENORI OSADA¹ KENYA SATO²

1. はじめに

1.1 背景

スマートフォンやタブレットが普及するにつれ、近年では拡張現実 (Augmented Reality:AR) の技術を使用したアプリケーションが見られるようになった。スマートフォンに付随するカメラを利用した AR アプリケーションや、スマートグラス等のウェアラブルデバイスを用いて AR 技術を活用するもの等がその例である。主な利用例としては、目的地までのナビゲーションシステム、AR マーカーを利用した観光地や施設の説明、スマートグラスを利用し、ディスプレイに必要な情報を表示させることで、医療現場や工場現場における作業のサポート等が挙げられる。

AR 技術を使用する利点は、現実空間とコンピュータ情報を同時に見ることができる点にある。現在車に搭載されているナビゲーションシステムの多くは、運転席の横にディスプレイを設け、経路情報等を表示させるのが一般的であり、現実空間である車前方向と、コンピュータ情報であるナビゲーション情報を同時に見ることは出来ない。医療現場や工場現場においても、資料に目を移しながらの作業は非効率である。これらの問題点は、スマートグラス等のウェアラブルデバイスを用いることで解決を図ることが

できる。また、今後はスマートグラスが普及するにつれ、ユーザがプライベートで AR デバイスを使用することも考えられる。

しかしながら、現在の AR デバイス、特にスマートグラスに関しては情報を表示するものという側面が強く、表示された情報に対しての操作に関しては未だ確立された手法は存在しない。Google 社のデバイスである Google Glass は音声認識によって操作を行うのに対して、EPSON 社の MOVERIO は付随の専用コントローラによって操作を行う等、デバイスによって操作方法が異なる場合がある。音声認識での操作では、テキスト入力を行いやすいがポインタの操作は不可能である。専用コントローラによる操作では、ポインタの操作は行きやすいが手が塞がってしまう問題がある。このように、既存手法ではそれぞれメリットが存在するが、同様にデメリットも存在している。そこで本研究では、スマートグラスの操作性において、既存手法における問題点の解決を図り、スマートグラスに表示された情報に対して効率的な操作を可能とするシステムを提案する。

1.2 目的

本研究の目的は、スマートグラス上に表示された情報に対する操作性を向上させることである。今後スマートグラスが広く普及した際、情報を見るだけでなく、それらの情

¹ 同志社大学大学院 理工学研究科 情報工学専攻

² 同志社大学大学院 理工学研究科

報を操作できるシステムが必要となる。現段階においても、音声認識やジェスチャー認識等によって AR 情報を操作することが可能であるが、従来のコンピュータのようにクリック動作やテキスト入力を行うことには難がある。

本研究における操作性の向上とは、従来のコンピュータに対するマウスのように、クリック動作やテキスト入力等、表示された情報に対して様々な操作をハンズフリーで行うことができる、ということの意味する。ここでのハンズフリーとは、操作する手に何も持つ必要がなく、手が塞がっていないことを意味する。これらの操作を可能とさせることで、情報の切り替えや選択、情報の書き込みや発信等をスマートグラス上で容易に行うことができる。また、スマートグラスの強みであるハンズフリーを生かすために、コントローラ等の手が塞がるデバイスは使用しない。このようにしてスマートグラスの操作性を向上させることが、本研究の目的である。

2. 関連技術

2.1 画像認識での操作

近年発売されているスマートグラスには、前方空間認識用のカメラが搭載されていることが多い。そのカメラを利用し、スマートグラス前方の手や指を検知することで、ジェスチャー操作等を実現して情報の操作を行う手法が存在する。加藤らは HandyAR を開発する研究において、カメラから得られた情報を元に画像処理を行って手形状を算出している [1]。この研究では、入力画像から HSV 色空間で肌色抽出をし、Dynamic Programming(DP) マッチング等を行い手形状を算出する。

また、指によるジェスチャー認識によってクリック動作を実現することもできる。杉浦らは画像処理による指の認識とクリック動作の認識を行うことで、ハンズフリーでのクリック動作を実現させる研究を行っている [2]。指先位置の検出は HSV 色空間上で肌色領域を抽出することで行う。また、クリック動作は指先が急減速する動作を検出することで判定を行う。

この手法によるメリットは、ハンズフリーで AR 情報の操作を行うことができる点である。指位置をトラッキングし、ジェスチャー機能を実現することで、指をマウスカーソルの様に扱うことも可能である。しかしデメリットとして、画像処理による計算量が多いことから、認識率の低下や反応速度が遅くなる問題点がある。

2.2 音声認識での操作

音声認識を用いて、特定の言葉を発することによってスマートグラスを操作する手法も存在する。Google 社が開発したスマートグラスである Google Glass はこの音声認識によって操作するものとなっている。"Ok,glass" と発言してから特定の言葉を発することにより、様々な機能を呼



図 1 EPSON 社の MOVERIO と付随の専用コントローラ

び出すことができる。

この手法のメリットは、ハンズフリーであること、そしてテキスト入力が行いやすい点である。Krevelen らの調査によると、AR 上でのテキスト入力は、今後音声認識が主流となってくると述べている [3]。しかし、音声認識にもデメリットは存在する。それは、操作についての幅が狭いということである。この手法では、あらかじめ決められた音声を発することで、決められた動作を行うという形になるため、限られた動作しか行うことが出来ない。また、画面上の位置を指定することも困難であるため、表示された情報を選択するという動作を行うことが難しい。

2.3 専用コントローラの利用

スマートグラスに付随する専用コントローラを使用し、画面上の操作を行う手法が存在する。EPSON 社が開発した MOVERIO にはタッチパッドとボタンが付いた小型専用コントローラがあり、タッチパッドを操作することでスマートグラスの画面上のポインタを動かすことができる。図 1 は、MOVERIO 本体と付随の専用コントローラである。

この手法でのメリットは、操作精度と反応速度が良い点である。コントローラを使用して操作を行うため、画像処理や音声処理と比較すると計算量が少なく、動作速度が早い。タッチパッドだけでなく、様々な物理ボタンを搭載することができるため、操作の幅が広いこともメリットの一つである。デメリットとしては、コントローラを使用するために手が塞がってしまう点が挙げられる。

2.4 指輪型ウェアラブルデバイスの利用

近年では、ジェスチャー操作を行うことに特化した指輪型ウェアラブルデバイスが開発されてきている。このデバイスには加速度センサ、ジャイロセンサ等のセンサが搭載されており、傾きや動きが検知できるものとなっている。このデバイスを指に装着し、その指でジェスチャーを行うことで各種センサがジェスチャーの種類を認識し、それに応じた操作ができる。

指輪型ウェアラブルデバイスはあくまで動作を検知する

ものであり、指位置やこのデバイスの位置を知ることは不可能である。このことから、スマートグラスの画面上の位置を指定することや、ポインタの操作をすることは困難である問題点がある。

2.5 既存手法における問題点

スマートグラスを操作する手法は様々であるが、それぞれに長所と短所が存在する。画像認識において画面上のポインタを移動させ、任意の位置を指定する操作を行う場合、クリック判定を行うために何らかのジェスチャーをしなければならず、判定に時間がかかってしまう問題がある。同様にドラッグ操作の判定に関しても、画像認識で行うことは難しい。指動作を検知する際、そのトラッキングの開始、終了判定を行うために、何らかのアクションが必要になってくるからである。コントローラを使用する際は物理ボタンを利用することでクリック動作を行えるが、画像認識の場合は、このワンアクションを入れる分の時間がネックになってしまう。

また、スマートグラス上でテキスト入力を行う手法に関しても、未だ確立された手法が無いのが現状である。音声認識でテキスト入力を行う手法があるが、公共の場や静かな場所において言葉を発するのは厳しく、使用場所が限定されてしまう問題がある。画面上に仮想キーボードを表示させ、ジェスチャーで文字を入力する手法も存在するが、前述の通り画像認識でのクリック判定が困難であり、入力速度に問題が生じる。コントローラを使用する場合は、スマートグラス本来の強みである、ハンズフリーという利点を消してしまう問題点がある。McCallらは、スマートグラス上でテキスト入力において既存システムを比較し評価を行ったが、全て一長一短であり現段階でスマートグラスに最適なテキスト入力の方法は無いと結論付けている [4]。

以上のことから、ハンズフリーで画面上のポインタを動かすことが可能で、クリックやドラッグの判定を物理コントローラに劣らない速度で行い、テキスト入力も可能とする新たな手法が、今後 AR 情報の操作に必要である。

3. 提案システム

3.1 概要

本提案システムは、ポインタ操作、クリックとドラッグの実現、テキスト入力をハンズフリーの既存手法と比較して素早く行うことを目的とする。スマートグラスに対する指の相対位置を求めることで、指をカーソルのように扱い、AR情報の操作性の向上を図る。指に赤外線 Light Emitting diode(LED)、スマートグラスに赤外線感知用の Complementary Metal Oxide Semiconductor(CMOS) センサを装着し、CMOS センサが赤外線光源をトラッキングすることで、指動作を認識する。また、指に装着する赤外線 LED に点灯、消灯を制御する物理ボタンを搭載する。これにより、

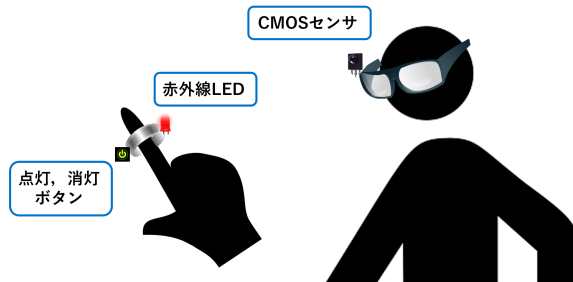


図 2 提案システム利用時の模式図

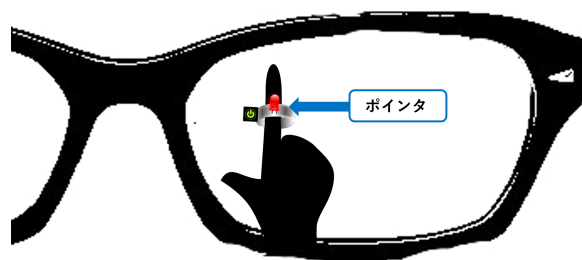


図 3 提案システム利用時の目線

赤外線が点灯している場合は指動作をトラッキングし、消灯している場合はトラッキングをしないという制御を行う。

図 2 は提案システム利用時の全体の模式図である。スマートグラスを装着し、前方で指を動かすことによって AR 情報の操作を行う。また、図 3 は操作している際の実際の見え方を示したものである。指に装着した赤外線 LED の位置と画面上のポインタの位置を同期することで、直感的な操作が可能である。

画像認識による手法では、手形状を算出するための画像処理における計算量の多さが問題となっていたが、赤外線のトラッキングに CMOS センサを利用することで、計算量を減らすことができる。計算量が減ることにより、システムへの負荷、動作速度が早くなり、AR 情報操作のレスポンスを向上させることができる。また、コントローラの使用と比較すると、ハンズフリーである点が利点として挙げられる。

3.2 システム構成

本提案システムでは、以下の機材を使用する。

- スマートグラス
- CMOS センサ
- 赤外線 LED
- 赤外線の点灯、消灯ボタン

スマートグラスに CMOS センサを装着する。CMOS センサは、スマートグラス前方の赤外線光源を感知できるものとする。人差し指に赤外線 LED を装着し、同じく人差し指の同じ位置に赤外線の点灯、消灯ボタンを装着する。

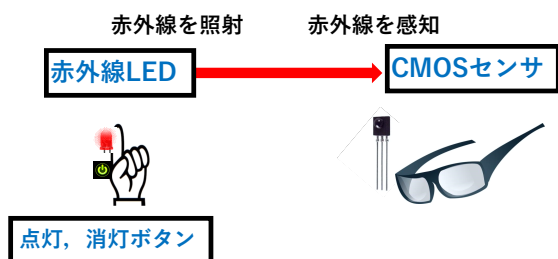


図 4 システム構成図

この際、赤外線 LED は手の甲側に付いているものとし、点灯、消灯ボタンは親指で押せる範囲に装着されているものとする。ボタンが押されていない状態では赤外線は消灯状態であり、ボタンを押している間だけ赤外線が点灯するものとする。図 4 は、提案システムの構成図である。スマートグラスに搭載された CMOS センサが前方の赤外線を検知し、赤外線光源の 2 次元座標を取得する。この赤外線光源をトラッキングすることにより、指動作を認識する。CMOS センサを常に動作させておくことで、赤外線が点灯している場合はその光源をトラッキングし、消灯している場合は赤外線を検知しないためにトラッキングをしない状態となる。

3.3 ポインタ操作

スマートグラスの画面上において任意の点を指定するのは、赤外線 LED を装着した指を動かすことによって行う。スマートグラスを装着した際に実際に見える指位置、つまり赤外線 LED の位置と画面上のポインタの位置を同期させることにより、赤外線を点灯させてトラッキングをせずともポインタの位置を知ることができる。これにより、カーソルを移動させる操作ではなく、指で絶対位置を指定することによる操作が可能となる。

3.4 指位置と画面上のポインタの同期

スマートグラスに搭載された CMOS センサが感知する赤外線光源の位置と、実際に人の目によって見える赤外線 LED の位置は、視差の関係で同位置には見えない問題がある。したがって、スマートグラスを装着した際に見える赤外線 LED の位置と、画面上のポインタの位置を同期させる処理が必要となる。図 5 と図 6 は、指を奥へ移動させた時の視線と CMOS センサとの視差を示したものである。実際には指をスマートグラスに対して横方向へは移動させていないものの、CMOS センサから見ると横方向へと移動しているものとして感知してしまう。

トラッキングする赤外線光源が 1 点のみであるため、

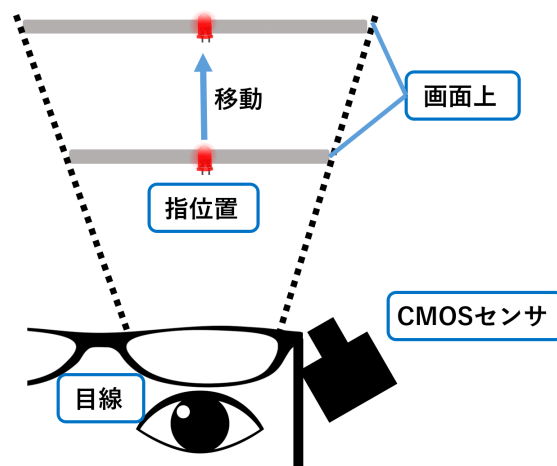


図 5 指を動かした際の実際の視線

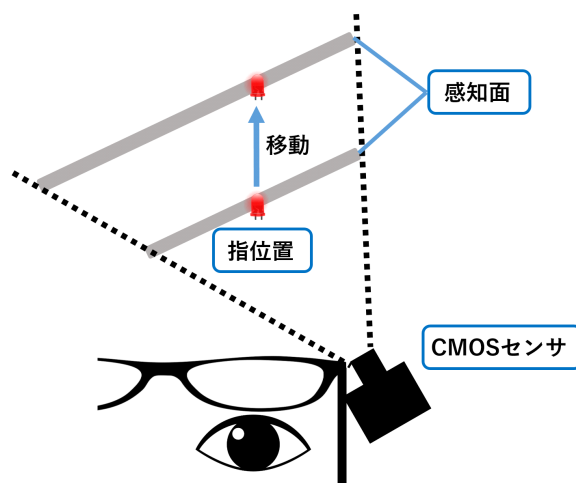


図 6 指を動かした際の CMOS センサでの感知

CMOS センサで得られる情報は 2 次元座標のみであり、奥行きを検知することが出来ない。したがって、各ユーザはあらかじめスマートグラスからどの程度離れた場所で指を操作するのかを指定する。その場所に仮想的な 2 次元平面を置き、その平面上で指を操作することで、指位置と画面上のポインタ位置を同期させる。

3.5 クリック判定

スマートグラスの画面上でのクリックは、赤外線 LED を点灯することによって行う。親指で赤外線 LED 点灯、消灯ボタンを押すことで赤外線が点灯し、消灯した段階でその地点においてクリックをしたと判定する。また、ボタンを押している間はドラッグの判定を行うものとする。

3.6 テキスト入力

スマートグラスの画面上に仮想キーボードを表示し、これを操作することによりテキスト入力を行う。また、キーボードを利用せず手書きによる描画によってテキストを表

現することもできる。この場合は、赤外線 LED が点灯している際に描画を行うことで、字の書き始めと書き終わりを判定する。

4. 実装

4.1 実装環境

スマートグラスは EPSON 社の MOVERIO を使用し、CMOS センサは Nintendo Wii リモコンを使用する。Nintendo Wii リモコンの先端部分には赤外線感知用の CMOS センサが搭載されており、赤外線光源をトラッキングすることが可能である。また、Bluetooth 接続に対応しており、同じく Bluetooth 接続に対応した外部機器へとセンサ情報を送信することができる。赤外線 LED は指に付ける事ができるものとし、また、赤外線の点灯、消灯を制御する物理ボタンを設け、このボタンを押している間は先端の赤外線 LED が点灯する仕組みとなっている。また、今回 Wii リモコンから得たセンサ情報を処理し、プログラムを動作させるコンピュータとして、Microsoft 社の Surface Pro 4 を使用する。使用 OS は Windows10 であり、使用プログラミング言語は Java である。

4.2 指動作の認識

Wii リモコンを MOVERIO の右側へ装着し、MOVERIO の前方で赤外線 LED を装着した指を動かすことで、この赤外線光源をトラッキングし、指動作の認識を行う。コンピュータと Wii リモコンを Bluetooth 接続することで、Wii リモコンが感知した赤外線光源の位置は 2 次元座標としてコンピュータへと送信され、この座標をトラッキングすることで指動作を認識する。

4.3 クリックの実装

本実装では、物理ボタンを押して赤外線 LED が点灯した時点で、クリックの判定を行う。赤外線 LED が点灯すると、Wii リモコンがその赤外線光源の 2 次元座標を取得し、コンピュータへと送信する。そして、コンピュータ上の Java プログラムがその地点でのクリック動作を行うことによって、赤外線 LED が点灯した地点でのクリックを実装する。

4.4 文字描画

物理ボタンを押して赤外線 LED が点灯した時点で文字の書き始めを判定し、赤外線 LED が消灯した時点で文字の書き終わりを判定する。Wii リモコンは赤外線を感知するとリアルタイムで 2 次元座標をコンピュータへと送信するため、指側において赤外線を点灯した状態で操作をすると、その操作のトラッキングをすることが可能である。線描画プログラムについては Java の API である Swing を使用し、Wii リモコンから受信した 2 次元座標の点を順に線

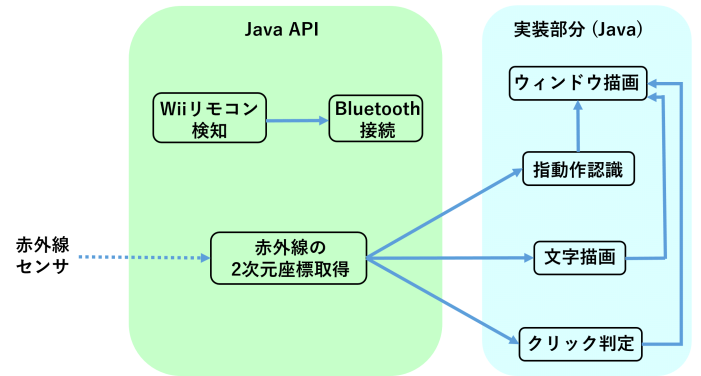


図 7 実装システムのモジュール構成

で結んでいくことにより、文字描画を実装する。これにより、物理ボタンを押した状態で指を動かすことで、線を描画することができる。

4.5 動作概要

図 7 は、実装システムにおけるモジュール構成である。Wii リモコンを制御する API として Java のライブラリである wiiusej を利用することで、コンピュータと Bluetooth 接続された Wii リモコンが感知したセンサ情報を取得する。CMOS センサから得られた赤外線光源の 2 次元座標を利用し、指動作認識、文字描画、クリック判定を行い、それぞれの処理結果をコンピュータ上のウィンドウへと表示する。

5. 考察

5.1 文字描画

画像認識による手法では、カメラで得た画像から手形状や指位置を検出する画像処理が必要であり、計算量が多くなってしまいう傾向がある。しかし、CMOS センサが赤外線を感知してこの光源をトラッキングする本提案手法では、画像認識による手法と比べて計算量が少ないという特徴がある。計算量の違いはシステム動作速度の違いとなり、ポインタの反応速度として表れてくるため、ポインタ操作の速さに関しても本提案システムが優位である。

5.2 操作性

本提案システムでは、ハンズフリーでポインタ操作、クリックやドラッグ操作、仮想キーボードを表示させることによるテキスト入力を行うことができる。専用コントローラを使用する場合でも同様の操作を行うことができるが、コントローラを持つ手が塞がってしまうという問題がある。画像認識による手法では、クリックやドラッグ動作を行うために何らかの動作が必要であり、その動作を判定する時間がネックになってしまう問題がある。音声認識による手法では、テキスト入力においては入力速度の点で他手法に

勝っているが、ポインタ操作が不可能である問題がある。指輪型デバイスを利用する場合でも、ジェスチャーでの操作は行えるが、ポインタ操作を行うことは困難である。

このことから、ハンズフリーでポインタ操作、クリックやドラッグ操作、テキスト入力を行える提案システムを利用することによって、スマートグラスの操作性が向上した。

6. まとめ

近年では AR 技術の普及によって、様々な場面においてスマートグラスが利用されている。しかしながら、現在のスマートグラスは AR 情報を見るデバイスという側面が強く、表示された AR 情報に対する操作方法は未だ確立した手法は存在しない。今後スマートグラスが仕事現場のみならず、私生活などのプライベートにおいても利用されることを考慮すると、スマートグラス上に表示された AR 情報に対して効率的に操作ができる手法が必要となってくる。

近年スマートグラス上の情報を操作する手法としては、画像認識を用いる手法や、専用コントローラを用いる手法など様々なものが存在する。それぞれの手法において利点があるが、同様に欠点も存在している。本研究では、既存手法に存在していた問題点を、赤外線を用いて動作をトラッキングすることにより解決した。

謝辞

本研究の一部は JSPS 科研費 (JP26540038) の助成を受けたものである。

参考文献

- [1] 加藤 喬, 近藤 裕介, 甲藤 二郎, "手をインターフェースとした拡張現実感システム HandyAR の開発", 電子情報通信学会技術研究報告, Vol.105, No.533, pp.13-18, 2006.
- [2] 杉浦 篤志, 豊浦 正広, 茅 暁陽, 西口 敏司, "拡張現実感のための直感的クリックインタフェース", 電子情報通信学会論文誌, Vol.97, No.9, pp.1426-1436, 2014.
- [3] D.W.F.van Krevelen, R.Poelman, "A Survey of Augmented Reality Technologies, Applications and Limitations", The International Journal of Virtual Reality, Vol.9, No.2, pp.1-20, 2010.
- [4] Roderick McCall, Benoit Martin, Andrei Popleteev, Nicolas Louveton, "Text entry on Smart Glasses", Human System Interactions (HSI), 8th International Conference, pp.195-200, 2015.