

拡張現実感における素手による仮想物体操作のための 隠面処理の検討

石津 貴弘^{a)}・坂本 眞人^{b)}

A Study on Hidden Surface Removal for Manipulation of Virtual Objects with Bare Hands in Augmented Reality

Takahiro ISHIZU, Makoto SAKAMOTO

Abstract

With the development of Augmented Reality (AR) technology in recent years, it has become possible for user to manipulate objects in virtual space as if they were operating objects in real space. In addition, it has been attracted on interaction technologies that allow users to manipulate virtual objects directly with their bare hands, aiming for more intuitive manipulation similar to real space. Therefore, we believe that a more accurate overlay technique between real and virtual space is required for a more seamless interaction. On the other hand, in augmented reality, a three-dimensional (3D) model is superimposed later on the image in real space, so that it is always displayed in front of the hand. Thus, in some cases, it becomes an unnatural scene (occlusion problem). Interaction manipulations in augmented reality will be increasingly required in the future, and more precise manipulations will be required. Therefore, we seek to consider hidden surface removal along the user's fingertips. In this study, the depth relationship between the user's hand and the virtual object is considered by acquiring depth information of the user's finger using a depth sensor. In addition, we propose a method to extract the area of the hand in real time by analyzing the state of the color distribution near the acquired finger position. Also, when performing an operation of grasping a virtual object, the thumb and other fingers could not be distinguished, and hidden surface processing was sometimes performed inappropriately. Therefore, in this study, we propose a method that uses the Canny method to distinguish between the thumb and other fingers and perform appropriate hidden surface processing. In this way, this system realizes the removal of hidden surfaces along the area of the user's hand. In the evaluation experiment, by using the hidden surface removal of this study, the boundaries of the fingers are distinguished and the outline of the fingers is clarified. Confirmed that it can be processed.

Keywords: Augmented reality (AR), Hidden surface removal, Virtual Object, Hand Detection

1. はじめに

拡張現実感 (Augmented Reality:AR) 技術は、我々が実空間から受ける知覚情報に、コンピュータによって生成された情報を重ね合わせる技術である。また、拡張現実感には教育、医学、娯楽などの幅広い分野で使用されており、その応用は多岐にわたる^{1),2),3)}。さらに、近年、拡張現実感技術はますます発展し、ユーザと仮想物体のインタラクションを実現するための研究も行われている^{4),5),6),7)}。

拡張現実感におけるインタラクションに関する従来研究⁸⁾では、ユーザにマーカや特殊な手袋などを置くことで、ユーザの位置情報を取得し、仮想物体との相互作用を実現した。ただし、これらの方法は、着用による体重などの不快感をユーザーに与える場合がある。

現実空間における実際の生活で、物体とのやり取りの主な手段は手であるため、拡張現実感におけるインターフェイスはユーザが素手で仮想物体を操作できる必要であると考えられる。

一方、拡張現実感で表示される3次元(3D)モデルは、実空間上の画像に後から重畳されるため、3Dモデルは常に前面に表示される。そのため、拡張現実感において、ユーザが素手で仮想物体とのインタラクション操作を行うとき、ユーザの手は仮想物体によって隠される。したがって、不自然な描写になり、ユーザは仮想物体と自分の手との前後関係を認識できず、仮想物体を操作することが難しくなる。

先行研究⁹⁾では、透明な3Dモデルを使用し、3Dモデルはユーザの各指を追跡させ、ユーザと3Dモデルの深度情報に基づいて隠面処理を行っていた。しかし、

a) 工学専攻機械・情報系コース大学院生

b) 情報システム工学科准教授

追従する 3D モデルが指よりも大きいため、実際の指よりも広い範囲が前面に表示されていた (図 1.1)。



図 1.1 : 先行研究における隠面処理

ここで、2020 年春から我が国において、第 5 世代移動通信システム (5G) サービスの商用化が始まる¹⁰⁾。

これまで 1G から 4G に至るまで、通信速度の向上が進んできた。5G もより高速化を実現するものであるが、5G ではさらに、「多数同時接続」、「超低遅延」といった特徴を持つ。

5G は来るべき IoT 時代の重要な基盤となるものとなり、拡張現実感と組み合わせることで、ユーザがリアルタイムにさまざまな情報をわかりやすく利用することに繋がると考えられる。

そのような背景を含めると、今後、拡張現実感におけるユーザと仮想物体とのリアルタイムに行うインタラクション操作に関して、ますます高い精度が求められると考えられ、拡張現実感におけるユーザと仮想物体とのリアルタイムに行うインタラクション操作に関して、ますます高い精度が求められると考えられる。

そこで、本研究においては、拡張現実感におけるインタラクション操作において、ユーザの手と仮想物体との隠面処理における見た目の部分に着目し、指先に沿った範囲での隠面処理を実現することにより、見た目の精度向上を目指す。

2. システムの構築

ユーザの手指の輪郭に沿って隠面除去を行うために、本システムは、ユーザの手指の深度情報、色情報およびエッジを取得することにより、ユーザの手領域を検出する。

まず、手領域を検出するために、本システムは色情報に基づいた RGB 画像を使用する。さらに、「肌色」領

域をリアルタイムで取得する。このときに使用する「肌色」は、リアルタイムにその色の範囲を更新し、定義する。そして、そこで定義された「肌色」の範囲内に入る色を手領域と認識する。

手指の深度情報の取得には、手指の深度情報の取得に特化したセンサである、Leap Motion Controller を使用する。取得した手指の深度情報に基づいて、親指に青色の点群を追従させ、他の指に緑色の点群を追従させる。

これらの青色および緑色の点群は、手が、Web カメラに向かって一番奥の仮想物体より前面にある場合のみ、その手指の位置に表示されるように、図 2.1 のように「境界線」を設定する。境界線の位置よりユーザの手指が、Web カメラから向かって手前かそうでないかにより、青色および緑色の点群を表示するか否かを制御する。つまり、境界線の位置よりも Web カメラから向かって手前側であれば、青色および緑色の点群を表示し、境界線の位置よりも Web から向かって奥側であれば、青色および緑色の点群を表示しないものとする。

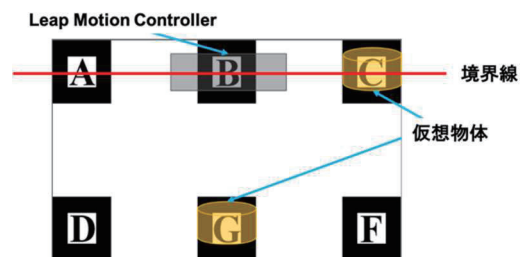


図 2.1 : システムの構成

また、仮想物体を複数表示する場合に対処するため、青色および緑色の点群と境界線より Web カメラから向かって手前側の仮想物体との関係を区別するために、Z バッファ法を用いる。

このようにして、ユーザの手が仮想物体よりもカメラから近い場合、ユーザの手は仮想物体の前面に正しく表示され、ユーザは自身の手と仮想物体との前後関係を把握できる。

指同士の区別は、親指のエッジ検出を行うことで、親指とそれ以外の指とを区別する。

このようにして、手の輪郭が抽出され、手の輪郭に沿った隠面処理を実現する。

ここで、親指とその他の指とで区別する理由としては、物体を掴む操作において、親指が仮想物体よりも手前に、その他の指が仮想物体よりも奥に位置する場面が多いと考えられるからである。

2.1 システム処理の流れ

拡張現実感では実空間の映像に対して仮想物体を後から重畳する。このため、仮想物体が実物体よりも奥に存在するときでも、その実物体に隠されることなく仮想物体が常に前にあるかのように表示される(図 2.2)。これは、拡張現実感におけるオクルージョン問題と呼ばれており、ユーザが仮想物体とのインタラクションを行う際に、仮想物体と手との前後関係を正しく認識できなくなる。このため、ユーザは視覚的に違和感を抱き、直感的な操作を行うことができなくなると考えられる。そこで、本研究では Leap Motion Controller から得られる手指の 3 次元座標と Web カメラから得られる RGB 画像を用いて、手と仮想物体の仮想空間上での隠面処理により、この問題を解決する。

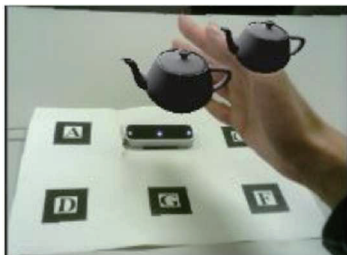


図 2.2 : 隠面処理をしない場合の AR 画像

以下にその手順を説明する。

1. Web カメラによる RGB 画像を取得し、肌色の領域を抽出することで手の領域を検出し、2 値画像(図 2.3)を生成する。
2. Leap Motion Controller により手の 3 次元座標を取得し、手のある部分が仮想物体よりも前にある場合、その部分をマスキングした画像(図 2.4, 2.5)を生成する。
3. 実空間と仮想物体を合成し、常に仮想物体が前面に表示される画像を生成する。

4. 手順 1. により抽出された手の領域にもとの RGB 画像を合成し、常に手が前面に表示される画像を生成する。

5. 手順 2. の画像におけるマスキングされた領域に手順 4. の画像を合成する。

以上の手順により、ユーザの手が仮想物体よりも手前側にある場合には、ユーザの手が仮想物体の手前側に正しく表示され、ユーザは自身の手と仮想物体との前後関係を把握することができる。

2.2 手領域を示す 2 値画像の生成

手領域を検出するため、手の「肌色」の領域の抽出を行う。「肌色」はリアルタイムにその範囲を決定する。次に、抽出した領域とそれ以外の領域で 2 値化を行い、膨張および収縮を 2 回ずつ行うことにより、ノイズを除去する。生成した画像を図 2.3 に示す。



図 2.3 : 2 値画像

2.3 深度情報を考慮した画像の生成

本システムは、Leap Motion Controller により取得した指の位置情報に基づいて、青色の点群(親指の位置に追従)と緑色の点群(親指以外の指の位置に追従)をプロットした画像を生成する。ユーザの親指が境界線より前面にあるとき、青色の点群は親指の各関節の位置に表示される(図 2.4)。

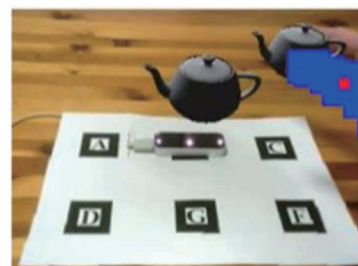


図 2.4 : 親指に追従する青色の点群

さらに、親指以外のユーザの手指が境界線より Web カメラから向かって前側にある場合、親指以外の指の各関節の位置に緑色の点群が表示される(図 2.5)。

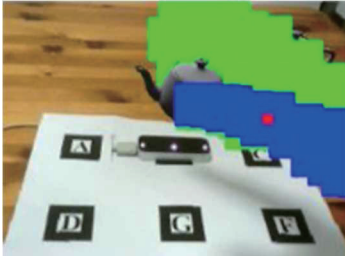


図 2.5 : 手全体に追従する緑色の点群

さらに、本システムにおいては、深度の異なる複数の仮想物体を処理できる。奥側の仮想物体とユーザの手との間との前後関係は、青色の点群と緑色の点群が表示されるかどうかによって判断できる。

一方、手前の仮想物体とユーザの手との間との前後関係は、Zバッファ法を使用して決定される。Zバッファ法を使用すると、手指が仮想物体の背後に位置すると、指の位置に続く青色の点群と緑色の点群が仮想物体によって隠される。

このようにして、本システムは、複数の仮想物体が存在する場合でも、ユーザの手と仮想物体との間の前後関係を判別できる。

2.4 手が前面に表示される画像の生成

図 4.5 における白い領域において、図 4.2 の画像を図 4.4 の画像に置き換えることで、手が常に仮想物体に対して前面に表示される画像(図 4.8)を生成する。

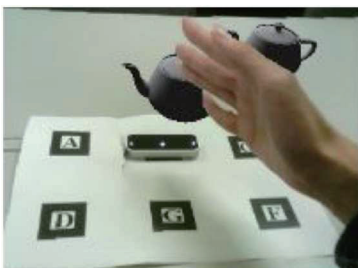


図 2.6 : 手が仮想物体の前面に表示される画像

2.5 結果画像の生成

図 2.2 の画像に対して、図 2.4 または図 2.5 の画像の緑色および青色の点群の領域だけを、図 2.6 の合成後の画像に置き換えることで、深度情報に応じて、仮想物体よりも前に手が表示される画像(図 2.7)を生成する。ユーザはこの画像を基にした映像を基に、仮想物体とのインタラクション操作を行う。

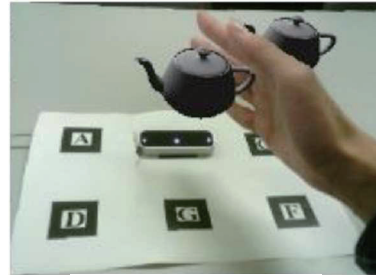


図 2.7 : 結果画像

3. 手領域の検出

本システムは、ユーザの手領域を検出するために「肌色」の領域を抽出する。

手領域は、肌の色や照明などの周囲の環境に左右されないようにすることが望ましい。よって、「肌色」の色の範囲はリアルタイムに決定することとする。

本研究のような、ヒューマン・コンピュータ・インターフェイス(HCI)システムにおいては、手領域を検出するための肌色検出のためのアルゴリズムが必要である。肌色検出のためのアルゴリズムには、2つの問題がある。1つは人間の肌の色調が可変的であること、もう1つは背景が複雑であることである。人間の肌の色調の変化は、同じ民族グループ内または異なる民族グループ内の人ごとに簡単に見ることができる。また、リアルタイムで使用される背景は、通常、照明条件が変化するため複雑になる。したがって、これらの条件下で適切に機能するには、優れた肌色の検出および認識アルゴリズムが必要である。

本稿では、背景の変化や照明条件の変化にリアルタイムに対応しながら、ユーザの肌色領域を検出する手法を提案する。これにより、手領域を適切に検出し、拡張現実感における隠面処理を、手の輪郭に沿った範囲で行うことが可能となる。

$$G_{xy} \in (0,255), B_{xy} \in (0,255), \text{if } xy \in C_{hand} \quad (3.1)$$

3.1 手領域候補の検出手法

本システムは、Leap Motion Controller から取得した位置情報により、親指の位置に青色の点群をプロット、その他の指および手の位置に緑色の点群をプロットする(図 2.4, 図 2.5).

しかし、ここで表示される青色および緑色の点群は、手領域を含むが、それ以外の領域も含んでいる. そこで、手領域のみを検出する必要がある.

本節では、青色の点群および緑色の点群の画素における RGB 値を元に主成分分析を用いることで、肌色らしさを表す指標軸が決まるかを確認し、手領域が持つ「肌色」を定義し、その領域を検出する手法を述べる.

3.2 主成分分析を用いた予備実験

主成分分析とは、多変量の値をできるだけ情報の損失なしに、少数の総合的指標に要約する手法であり、比較するデータの数を削減できる.そのため、高速処理が実現でき、リアルタイム性が要求される場面に適していると考ええる.

本節では、まず予備実験として、図 2.5 の青色の点群および緑色の点群の画素において、手領域を赤、それ以外の領域を青とする点をプロットする(図 3.1).

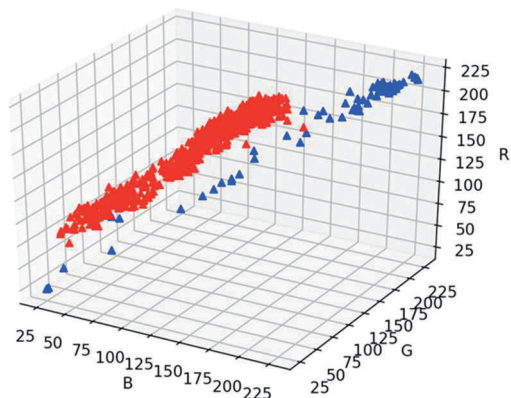


図 3.1 : 青色および緑色の点群における画素の分布

画像空間上の座標 (x, y) をもつ画素 xy の画素値 x_{xy} を縦に並べたデータ行列 x_D を式 5.1 のように定義する.

$$X_D = \{x_{xy} | x_{xy} = (R_{xy}, G_{xy}, B_{xy}), R_{xy} \in (0,255),$$

ここで、

C_{hand} : 図 2.5 の青色の点群および緑色の点群の画素の集合.

x_{mean} : 図 2.5 の青色の点群および緑色の点群の画素全体の平均値ベクトル.

表 3.1 : 固有ベクトル

	R	G	B
第 1 主成分	-0.575	-0.583	-0.575
第 2 主成分	-0.716	0.017	0.698
第 3 主成分	-0.426	0.813	-0.396

N : 図 2.5 の青色の点群および緑色の点群の画素数.

とする.

次に、 X_D を標準化するため、 X_D における平均を μ 、標準偏差を σ とすると、 $X_S = (X_D - \mu) / \sigma$ により変換する.

以下、変換後の X_D を X とする.

データ行列 X から、各軸間の関係を圧縮し、 Z の分散が最大になるように

$$Z = \omega_1 X_R + \omega_2 X_G + \omega_3 X_B \quad (3.2)$$

を定める.

ただし、 $\sum \omega_i^2 = 1$ とする.

式 3.2 および固有ベクトル(表 3.1)から、主成分得点を求める式は、

$$Z_{1xy} = -0.575R_{xy} + (-0.583)G_{xy} + (-0.575)B_{xy} \quad (3.3)$$

$$Z_{2xy} = -0.716R_{xy} + 0.071G_{xy} + 0.698B_{xy} \quad (3.4)$$

$$Z_{3xy} = -0.426R_{xy} + 0.813G_{xy} + (-0.396)B_{xy} \quad (3.5)$$

のように表される.

これらの式を用いて主成分得点を求め、その画素の分布を図 3.2 に示す.

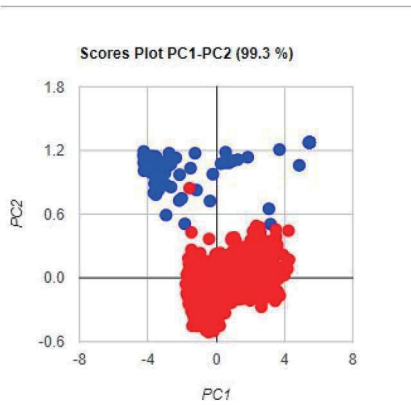


図 3.2 : 青色および緑色の点群における画素の分布

本研究では、主成分得点に閾値を定めることによって手領域とそれ以外の領域とを分離する。そこで、分布の状態と寄与率から、閾値の設定が容易と考えられる第1および第3主成分得点を用いて閾値を設定することとする。閾値は、箱ひげ図の要領で定める。ある一定の割合に入る RGB 値をもつ画素を肌色領域とみなすこととし、上限および下限を適切に定め、それぞれ閾値とする(図 3.3)。この時、手領域とその他の領域との区別を割合的に求めるため、手領域の重みを大きくする。これは、青色および緑色の点群の中心付近の重みを大きくすることで得られる。

図 3.3 のように、肌色領域とそれ以外の色領域とを分離するように、仮に閾値として $-2 < Z_1 < 4$, $-0.6 < Z_2 < 0.5$ と定める。

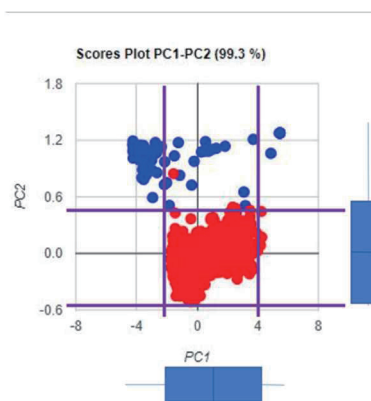


図 3.3 : 仮の閾値の設定

$-2 < Z_1 < 4$, $-0.6 < Z_2 < 0.5$ を満たす画素を白、それ以外を黒とする 2 値画像を生成する (図 3.4)。

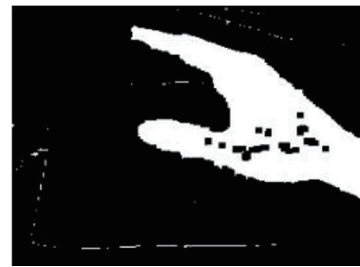


図 3.4 : 予備実験における 2 値画像

3.3 生成した 2 値画像におけるノイズの除去

本システムは、「肌色」を使用してユーザの手領域を検出する 2 値画像を生成するが、この 2 値画像には、ノイズが含まれている。

この 2 値画像は後の処理で使用されるため、ノイズが最終結果画像に反映されてしまう。

したがって、本節ではそのノイズを除去する手法について説明する。

まず、8 近傍で収縮と膨張を 2 回実行する。ただし、これだけではノイズを十分に除去できない。これは、大きなノイズを除去するために、膨張と収縮を複数回行う必要があるため、計算量が大きくなり、リアルタイム性が損なわれるということがあげられる。そのため、大きなノイズに関しては、次の処理を行うことで手領域のノイズを除去する (図 3.5)。

1. 赤い点(カメラに最も近い指に置かれる)を含む白い領域を手領域として決定する。
2. 手の内側を黒で塗りつぶし、手の外側を白で塗りつぶす。
3. 白と黒を反転する。これにより、手領域内のすべてのノイズが除去される。

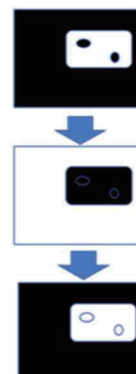


図 3.5 : ノイズ除去の手順

4. 指同士の区別

図 4.1 の白色の領域は手の領域を示している。ただし、この画像からは指同士の区別が明示されていない。



図 4.1：指同士の区別が明示的でない 2 値画像

よって、この画像に基づいて隠面消去を行うと、図 6.2 のように親指だけでなく、他の指巻き込まれて隠面処理されてしまう。

したがって、この問題を解決するために、エッジ検出を用いて親指と他の指とを区別する。



図 4.2：他の指を巻き込む

4.1 Canny 法によるエッジ検出

本研究では、エッジ検出の手法として、計算コストが低く、リアルタイムでの処理に適している Canny 法¹⁾を採用する。

Canny 法では、画像の輝度チャンネルに大きな変化がある部分をエッジ強度として取得する。強閾値と弱閾値 2 つの閾値が設定され、エッジ強度が強閾値を上回っている場合と、エッジ強度が弱閾値以上かつ強閾値未満ですでに輪郭と判定された画素に隣接している場合に輪郭の画素であると判定する。

さらに、エッジ検出時に抽出される親指の領域は、後の処理でエッジに囲まれた領域で埋められるため、エッジ検出時には、エッジは 1 つの明確なラインでなければならない。そのため、本研究では Canny 法を用いた、ユーザの親指のエッジを検出する手法を提案する。

4.2 親指領域の検出

親指の輪郭を検出するために、図 4.3 のように、赤色の点群から親指の付け根方向に向けて探索を行い、Canny 法で検出したエッジに当たった点があれば、そこに親指の輪郭があるとする。Canny 法で検出したエッジに当たった点がなければ、角度を変えて再び探索する。

このようにして、親指の輪郭を探索し、そこから少し進んだ位置を三角形の 2 つの頂点とする。また、親指の付け根から少し進んだ位置をもう 1 つの頂点とする。

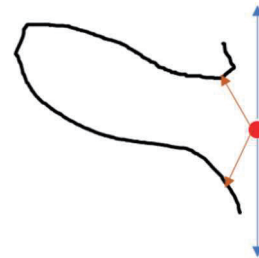


図 4.3：親指の輪郭検出

このようにして得られた 3 つの頂点をもつ三角形を図 4.4 に示す。

親指を検出するために、親指を三角形で囲み、親指で囲まれた内側の領域とエッジ検出で表示された境界を抽出する。この領域を親指領域と見なす(図 4.4)。

図 4.4 と深度を考慮した画像を合わせて結果画像を生成する(図 4.5)。

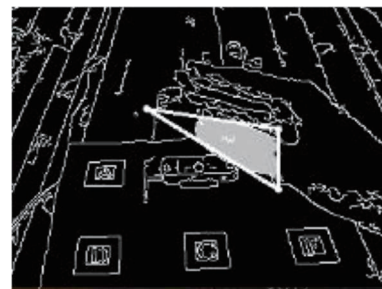


図 4.4：親指領域を明示的に表示

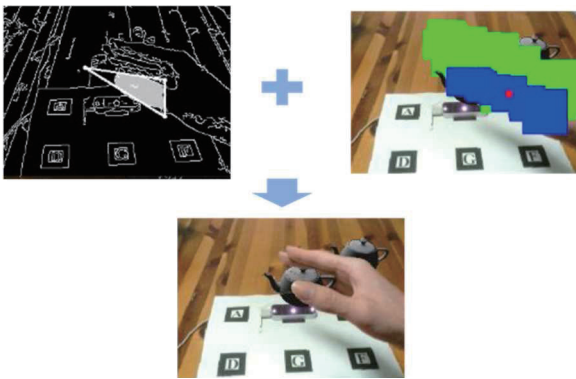


図 4.5 : 親指領域を明示的に表示することで得られた結果画像

5. 評価実験

本システムにおいて、肌色検出により、手指の領域検出の精度を検討する実験を行った。

320×240 pixel の画像を作成し、本手法における隠面処理を行う。このとき、手領域である画素を正しく判定した画素数、誤って手領域と判定した画素数、手領域であるが手領域と判定されなかった画素数および判定の精度をそれぞれ求め、その結果を表 5.1 に示す。

検出の精度に関して、さらに向上させるために、肌色の閾値が適切に設定できるようにする必要があると考える。

表 5.1 : 手領域の検出精度

	1人目	2人目	3人目
正しく判定 [pixel]	494	473	410
誤って手領域と判定 [pixel]	36	63	4
手領域と判定されない [pixel]	100	58	69
精度 [%]	78	80	85

6. おわりに

拡張現実感近年、教育、医療、エンターテインメントなど、幅広い分野に適用されている。将来的には、拡張現実感でのインタラクション操作のさらなる開発が期待され、より詳細な仮想物体の操作が必要になると考えられる。

したがって、本研究では、拡張現実感でのインタラ

クション操作に関するより詳細な作業に適用できることを目指して、指に沿った隠面除去を提案し、実装した。

本研究では、ユーザの指に沿った隠面除去を実現するために、手の輪郭の検出に注意を払った。そのため、本研究においては、手領域の検出のために「肌色」を主成分分析により決定、および抽出する手法を提案した。さらに、親指と他の指との区別がないため、不適切に隠面処理されてしまう問題を解決するため、本研究では、Canny 法を用いた親指のエッジを検出する手法を提案した。

評価実験では、本システムにより、リアルタイムに手領域の範囲を適切に決定することができ、特定個人に依存せずに、拡張現実感における指先に沿った隠面処理が実現できることが確認できた。

今後の課題として、本システムにおける処理の安定性の向上が挙げられる。また、肌色の検出について、さらなる精度の向上のため、肌色の閾値に関するパラメータを適切に設定する必要があると考えられる。

参考文献

- 1) KangdonLee, “Augmented Reality in Education and Training”, TechTrends, vol.56, no.2, pp.13-21, 2012.
- 2) Carmigniani J, FurhtB, Anisetti M, Ceravolo P, Damiani E, Ivkovic M, “Augmented reality technologies, systems and applications”, Multimedia Tools and Applications 51:341-77, 2011.
- 3) Abdelkader Bellarbi, Christophe Domingues, Samir Otmame, Samir Benbelkacem and Alain Dinis, “Underwater augmented reality game using the DOLPHYN”, Proceedings of the 18th ACM symposium on Virtual reality software and technology, ACM, 2012.
- 4) Takahiro Ishizu, Makoto Sakamoto, Takaaki Toyota, Masamichi Hori, Satoshi Ikeda and Amane Takei, “Proposal for Interaction Techniques for Intuitive Virtual Objects Manipulation in Augmented Reality”, Proceedings of the 2019 International Conference on Artificial Life and Robotics, pp.237-240, 2019.
- 5) Makoto Sakamoto, Takahiro Ishizu, Takaaki Toyota, Masamichi Hori, Satoshi Ikeda, Amane Takei, Takao Ito, Yu-an Zhang, “Proposal for Interaction Techniques for Intuitive Virtual Objects

Manipulation in Augmented Reality”, Journal for ROBOTICS, NETWORKING AND ARTIFICIAL LIFE, 2020.

6) 鈴木遼人, 鈴木奏太, 佐藤美恵. “素手での掴み動作に着目した AR システムの 視覚的改良”. 映像情報メディア学会誌, 2014, Vol 68, No. 7, pp. J299-J301.

7) 加藤昇, 石垣早都, 鈴木奏太, 佐藤美恵. “素手での掴み動作における仮想物体 の柔らかさ提示に関する検討”. 映像情報メディア学会誌, 2017, Vol 71, No. 3, pp. J121-J124.

8) Volkert Buchmann, et. al. “FingARtips-Gesture Based Direct Manipulation in Augmented Reality”. 2004.

9) R. Katahira, M. Soga. “Development and Evaluation of a System for AR enabling Realistic Display of Gripping Motions using Leap Motion Controller”, Procedia Computer Science 60:1595-1603, 2015.

10) “第 1 部 特集 人口減少時代の ICT による持続的成長”, 総務省. [online]<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/html/nd133420.html> (2020 年 1 月 25 日アクセス).

11) J. Canny. “A Computational Approach to Edge Detection”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol.8, no 6, pp. 679-698, 1986.