

視点と手とディスプレイの相対位置計測による 拡張現実感の自己身体融合感の向上

三留 愛美^{†1} 金子 直史^{†1} 齋藤 友彦^{†2} 鷲見 和彦^{†2}

概要: 拡張現実感において、画面外の自己身体と仮想物体とが融合した感覚を持つことは難しい。本研究では仮想物体に画面外から手を伸ばして触るというシナリオにおいて、触れている感覚が向上する手法を提案する。利用者の視点位置と手の位置をそれぞれ専用の 3D センサで検出し、位置情報に基づいて 3D ディスプレイ内の仮想物体を位置合わせすることにより、現実空間の手に接触したように見える仮想物体を表示する。この実現には、3D センサ・視点・ディスプレイの間で座標系の正確な校正や、3D 眼鏡を装着した利用者の視点の 3 次元位置・姿勢を安定して検出することが必要で、校正法と視点検出法についても言及する。最後に実験によりシステムの評価・考察を行う。

Strengthening of Fusion of Human Bodies and Augmented Reality by Relative Position Measurement of View Positions, Hand and Display

MANAMI MITOME^{†1} NAOSHI KANEKO^{†1}
TOMOHIKO SAITO^{†2} KAZUHIKO SUMI^{†2}

Abstract: In augmented reality (AR), it is a problem to obtain smooth registration between off-screen real human body and virtual objects on screen. In this research, we cope with the problem when an AR user touches a virtual object from outside of the screen. Our proposed method deploys conventional 3D display with liquid crystal 3D glasses. We introduced two 3D gesture sensors measuring the user's eye location and user's hands posture respectively. First, we measure user's hand position with a 3D hand gesture sensor and user's eye location with another 3D (RGBD) camera. The virtual objects are displayed onto the screen according to the user's eye position. The user's hand and the object interacts when they are making a contact. To realize this system, it is important to calibrate the two 3D sensors and the display, as well as reliable and stable detection of the user's eye position under the 3D glasses. In this paper, we describe the calibration process and the eye position detection in details. Finally, we describe the precision evaluation and user feeling evaluation of the experimental system.

1. はじめに

実世界の映像に仮想物体の映像を重ねる拡張現実感(AR: Augmented Reality)は、作業指示・案内・遠隔作業支援などへの応用が期待されている。多くの AR 応用では利用者は現実世界をコンピュータ画面の中で見るため、画面外の自己身体と仮想物体との重ね合わせの実現は難しい。例えば加藤らの Handy-AR[1]は画像から手を検出し、手と仮想物体の映像とを融合したが、利用者は画面中の自分の手を見ており現実感の提示が不十分である。また、Hoshi らの Touchable Holography[2]は表示に実空間に立体表示可能なホログラフィと超音波力覚デバイスを用いて実物の手の上に立体表示と触感とを実現したが、ホログラフィ装置や超音波アレイスピーカという特殊装置が必要で、表示が半透明なため完全ではない。

本研究では仮想物体に触っているように見える AR システムを安価に実現する手法を提案する。利用者の視点位置と手の位置をそれぞれ専用の 3D センサで検出し、その情

報と液晶シャッター眼鏡を使うタイプの 3D ディスプレイの映像との位置合わせにより、現実空間の手の上に仮想物体を 3D で重ねて表示する。この実現には、3D センサ・視点・ディスプレイの間で座標系の正確な校正や、3D 眼鏡を装着した利用者の視点の 3 次元位置・姿勢を安定して検出することが必要で、校正法と視点検出法についても言及する。最後に実験によりシステムの評価・考察を行う。

2. 既存技術の説明

従来から利用されている AR と 3D ディスプレイ表示のための座標変換について説明する。表示物体ごとに定義されたオブジェクト座標系(物体ごとの頂点などを指定するための座標系)での座標 $X = [x \ y \ z \ 1]^t$ を式(1)で画像表示座標系(画面左下を原点、画面の幅・高さを画素数とする座標系)での座標 $X' = [x' \ y' \ z' \ 1]^t$ に変換し表示する。

$$X' = M_{nd} M_{cn} M_{ec} M_{we} M_{lw} X. \quad (1)$$

但し、 M_{lw} はオブジェクト座標系から世界座標系(物体を置く 3DCG 空間全体の基準となる座標系であり、本研究では 3D ディスプレイの中心を世界座標系の原点としている)での座標 W への変換、 M_{we} を W から視点座標系(視点を

^{†1} 青山学院大学大学院理工学研究科
Aoyama Gakuin University Graduate School of Science and Engineering
^{†2} 青山学院大学
Aoyama Gakuin University

原点とした座標系)での座標 V への変換, M_{ec} は V からクリップ座標系(弱透視投影で得られる座標系)での座標 C への変換, M_{cn} は C から正規化デバイス座標系(画面中心を原点, 画面幅・高さ・奥行きを $[-1, 1]$ で表す座標系)での座標 D への変換, M_{nd} は D から画像表示座標系での座標 X' への変換である. 左右の目に対してこの座標変換を行うことで, 立体視用の画像を生成することができる. 図 1 は片方の目に対して仮想物体を表示するための座標系である.

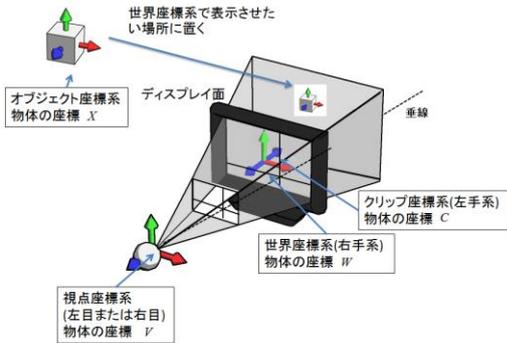


図 1 仮想物体表示のための座標系

Figure 1 Coordinate system for display virtual object.

3. 提案システム

3.1 システム概要

図 2 は本システムの構成で, 従来の 3D 眼鏡を使用する 3D ディスプレイを利用している. さらに, ユーザの視点・手を検出して仮想物体が手に接触して見えるよう表示するために, 視点検出用の RGB-D センサと手検出用の D センサを配置する. 視点を検出し易くするために 3D 眼鏡にカラーマーカを装着しておく. また, 図 3 が本システムを実行した時の様子である. 3D 眼鏡を掛けたユーザからは, 3D ディスプレイに 2 重に表示された仮想物体が, 立体的に, 実空間の手のひらにあるように見える.

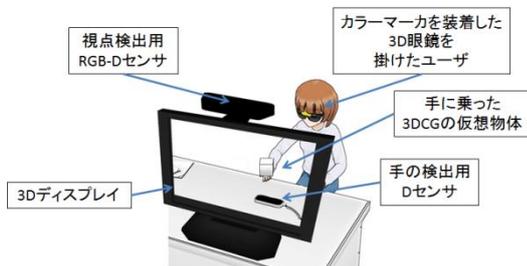


図 2 システムの構成

Figure 2 Configuration of our system.

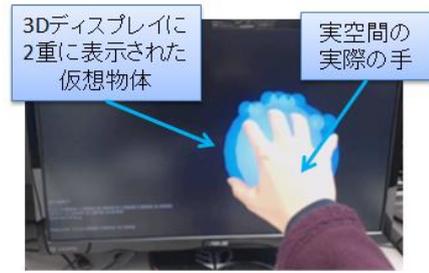


図 3 実行の様子

Figure 3 Execution example.

3.2 本システムの座標変換

本システムの仮想物体を表示するための座標変換について述べる. まず, マーカを RGB-D センサで撮影し, CamShift 法[3]で認識させる. 認識させたマーカの位置から, カメラから見た左目または右目の 3 次元座標 $E_k = [e_x \ e_y \ e_z]^t$ を算出する (3.2 節参照). 次に, D センサでセンサを原点とした手の 3 次元座標 $H_l = [h_x \ h_y \ h_z]^t$ を検出する. E_k, H_l を世界座標系での位置 $E_w = [e_x' \ e_y' \ e_z']^t$, $H_w = [h_x' \ h_y' \ h_z']^t$ に変換する (3.3 節参照). H_w を式 (1) の M_{lw} に, E_w と H_w を M_{we} と M_{ec} に適用する. オブジェクト座標 X に対して, 式 (1) の変換を行い, 立体画像を表示する. 図 4 が本システムの座標系である.

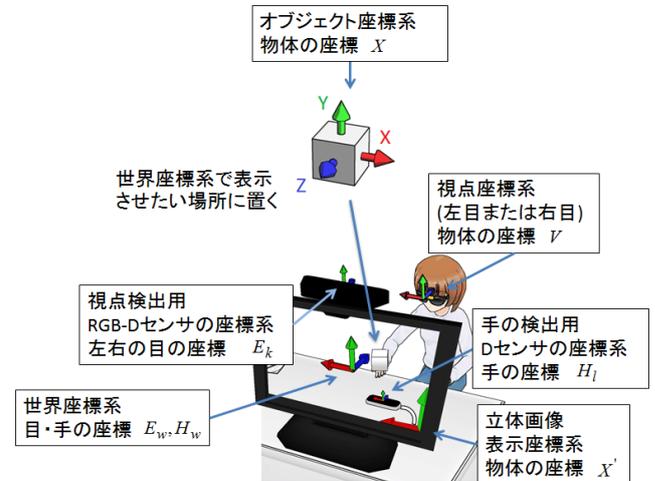


図 4 システムの座標系

Figure 4 Coordinate system of our system.

3.3 視点位置の検出

3D 眼鏡を掛けた状態で目の位置を検出するのは困難なため, 3D 眼鏡にカラーマーカを付け, CamShift 法で追跡. そのマーカから, 視点の大体の位置を計算する (図 5). 両目の間隔は約 6cm[4]とし, マーカの長さもこれに合わせる.

位置の計算は, センサの RGB 画像上でカラーマーカ部分を認識し, CamShift 法で認識された四角形範囲の中心座標 $X_M = [c_x, c_y, 1]^t$, 幅 w_m (四角形の短辺, ピクセル単位),

高さ h_m (ピクセル単位)と傾き θ から, 2次元アフィン変換で左右それぞれの目の位置をピクセル単位で求める. 認識範囲の中心座標 X_M を原点としたマーカ座標系での左右の目のピクセル位置 $P_l = [p_x^l \ p_y^l]^t$, $P_r = [p_x^r \ p_y^r]^t$ を,
 $\hat{p}_x^l = -h_m/2$, $\hat{p}_y^l = a \cdot w_m/2$, $\hat{p}_x^r = h_m/2$, $\hat{p}_y^r = a \cdot w_m/2$ とする. p_y については 3D 眼鏡に付けたマーカの短辺の長さの半分の $a=4.5$ 倍の位置を視点としている. P_l , P_r を 2次元アフィン変換 (θ で回転, X_M で並進) して, 画像の左上を原点とした画像座標系の左右の目の位置のピクセル値 $P_l' = [p_x^{l'} \ p_y^{l'}]^t$, $P_r' = [p_x^{r'} \ p_y^{r'}]^t$ を求める. この値は, フレームごとに誤差が出てくるため, 検出される位置が安定しない. そこで, カルマンフィルター[5]で視点の位置を推定する.

次に, そのピクセル値 P_l' , P_r' に対応するデプス情報(センサからの距離)を取得させるが, その部分のデプスが取れない場合があるので, その周辺部分のデプス値も取得し, 取れた値の平均をデプスの値とする. 得られたピクセル位置とデプス値からセンサ座標系での左右それぞれの目の3次元座標 E_k を取得する.

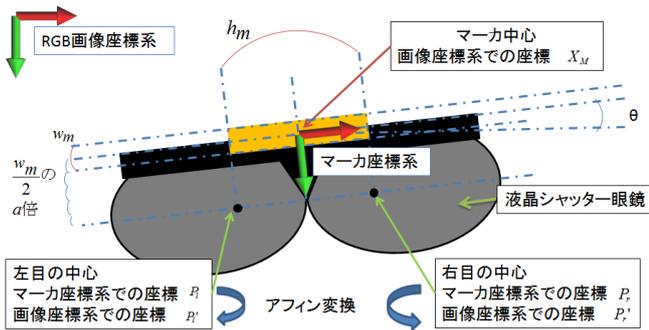


図 5 カメラから見たマーカと目の位置
 Figure 5 Positions of the marker and eyes when viewed from a camera.

3.4 座標系の校正 (キャリブレーション)

E_k , H_l を世界座標系での位置 E_w , H_w に変換するための行列 M_k , M_l を求めるための方法について述べる. 行列 M_k はまず, RGB-D センサでチェックボードのコーナー点を検出し(図 6), その点の座標を横に並べた行列を B_k とする. それらの世界座標系での対応点を実測し, それを A_k とする. $A_k = M_k B_k$ となる行列 M_k を求める. 行列 M_l は, 図 7 のように, 世界座標系の 3 軸とセンサの座標系の 3 軸がそれぞれ平行になるように設置し, x 軸の原点を揃え, ディスプレイからの距離 $l = [0 \ l_y \ l_z \ 1]^t$ を測り, これを並進行列にあてはめ M_l とする.

上で求めた M_k, M_l を使い, 目, 手の世界座標 $E_w = M_k E_l$, $H_w = M_l H_l$ を求める.

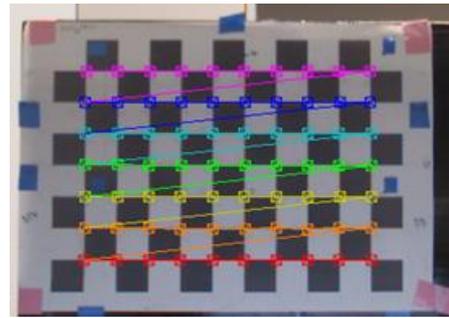


図 6 チェックボードのコーナー一点検出
 Figure 6 Corner detection of checkerboard.

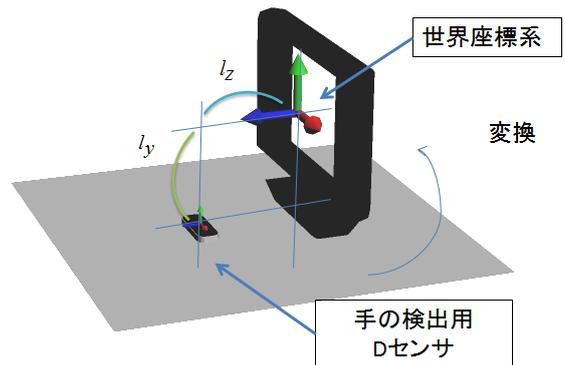


図 7 センサとディスプレイの位置関係
 Figure 7 Positions between the sensor and the display.

4. 評価実験・考察

本システムの評価のために, センサの計測誤差評価, 被験者によるアンケート評価, 本システムの誤差評価の 3 つの実験を行った.

4.1 実験 1 センサの計測誤差評価

本システムの実装では, 目の検出用 RGB-D センサとして Microsoft 社の Kinect, 手の検出用センサとして Leap Motion 社の Leap Motion を使用した. それぞれのセンサの計測誤差を求める. まず, センサの計測値と, センサの原点から実測した値を真値とした場合の計測値と真値の差を座標変換前の誤差とする. 次に, センサの計測値を行列 M_k, M_l によって世界座標系に変換した後の値と, 世界座標系の原点から実測した値を真値とした場合の計測値と真値の差を座標変換後の誤差とする. 座標変換前ではセンサ座標系の xyz 軸それぞれの真値とセンサ計測値の差, 座標変換後では世界座標系の xyz 軸それぞれの真値とセンサ計測値の差を表 1 に示す. 目検出用センサの認識では, 校正による座標変換によっていくつかの軸は誤差を小さくすることができ, 変換行列は有効であったと言える. 手検出用センサの認識には, まだ誤差があるので, その誤差を小さくできるように校正を行い, 並進のみでない座標変換をする

べきである。

表 1 目手検出誤差 (単位:cm)

Table 1 Error of eyes and hand detection.

	目検出誤差			手検出誤差		
	x 軸	y 軸	z 軸	x 軸	y 軸	z 軸
座標変換前	0.59	3.32	0.69	0.19	0.90	0.16
座標変換後	0.44	0.93	2.58	0.19	0.90	0.16

4.2 実験 3 システムの誤差評価

本システムの誤差を評価する。仮想物体を固定した位置に表示させ、図 8 のように指で仮想物体に触れようとし、触れたように見えたときにキーボードを入力し、指の 3 次元位置をセンサで取得させる。また、仮想物体の表示位置を真値とし、計測位置との差を誤差とする。仮想物体の左右から触れた場合を x 軸，上下から触れた場合を y 軸，手前から触れた場合を z 軸上の位置とし，xyz 軸それぞれの計測位置と真値の差を表 2 に示す。それぞれ 1cm 程度になり，本システムの誤差は 1cm 程度であると言える。

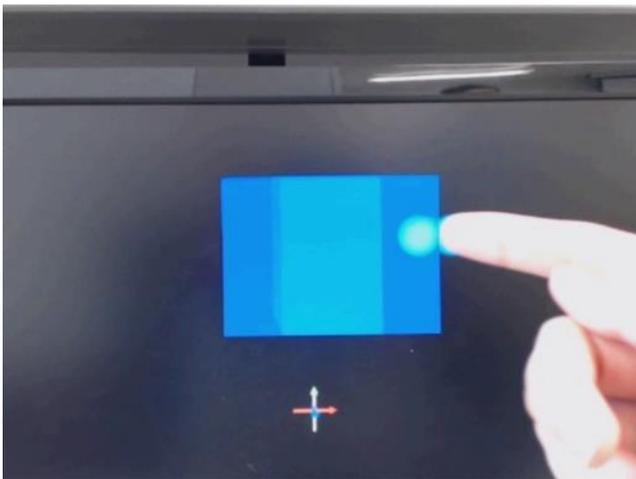


図 8 システムの誤差評価実験

Figure 8 Evaluation experiment of error of our system.

表 2 システムの誤差 (単位:cm)

Table 2 Error of our system.

	x 軸	y 軸	z 軸
誤差	0.86	1.10	1.01

4.3 実験 2 被験者による実験

本システムの現実感を被験者 5 人によるアンケートにより 5 段階で評価する。目を固定し，手を上下に動かしたとき，仮想物体が手に乗っているように見えるか(手乗り感)を現実感とする。手乗り感の評価を図 9 に示す。手を下に移動すると，下にずれて見え，手に乗っているように感じられないが，これは本来仮想物体で手が隠れるべき位置で

あり，投影面が手の向こうにあるので，手と物体の位置が逆転してしまうからだと考えられる。手を視点と同じ位の高さに上げると投影面が手で隠れないので手に乗っているように見せることが可能となった。よって本システムは，手で手前から物体を覆うように触る AR システムに適していると言える。

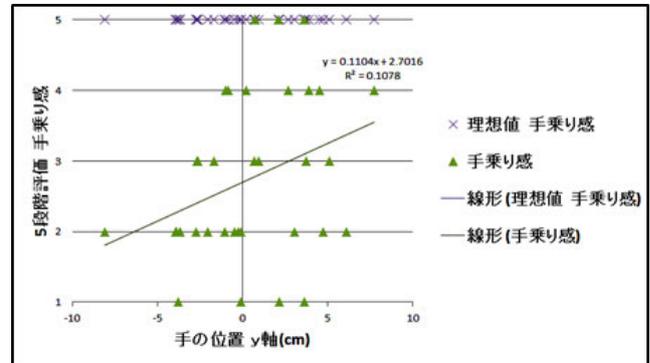


図 9 手を上下に動かしたときのずれの感覚

Figure 9 Feeling of gap

when you move the hand up and down.

5. まとめ

本研究では，3D ディスプレイと複数のセンサを用いて，仮想物体に触っているように見える AR システムを安価に実現する手法を提案した。これを実験により評価し，その有効性を示した。AR において，特殊なデバイスを使わなくても，仮想物体と自己身体の融合感を強化し，現実空間の手のひらで扱っているように見せることが可能となった。今後は，さらなる精度向上や視点位置検出の安定化をし，本システムを有効に使えるアプリケーション等の開発を行いたい。

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.535，pp.13-18 (2006).
- [2] T. Hoshi, M. Takahashi, K. Nakatsuma, and H. Shinoda: Touchable Holography, SIGGRAPH 2009, Emerging Technologies (2009).
- [3] Gary R. Bradski, Microcomputer Research Lab, Santa Clara, CA, Intel Corporation: Computer Vision Face Tracking For Use in a Perceptual User Interface, Intel Technology Journal, no.2nd Quarter, p. 15(1998).
- [4] 小曳尚，繁樹博昭，北崎充晃: パーチャルリアリティ空間における視聴覚の同時知覚：両眼視差と運動視差の効果(複合現実感, 仮想都市, 及び一般), 電子情報通信学会技術研究報告, Vol.106, No.234, pp.73-77 (2006).
- [5] R. E. Kalman. 1960. "A New Approach to Linear Filtering and Prediction Problems," Transaction of the ASME?Journal of Basic Engineering, pp. 35-45 (March 1960).