

# Kinect を用いた手指画像による入力システム の構築に関する基礎的検討

平成 28 年 3 月

日本大学理工学部電子情報工学科

中村高橋研究室

No.2033 経 乾

# 目次

1.はじめに .....	2
2.Kinect の概要 .....	2
2.1 Kinect v1 と Kinect v2 の区別 .....	2
2.1.1 Kinect の仕組みと原理 .....	2
2.1.2 Kinect の仕様 .....	3
2.1.3 Kinect の開発環境 .....	3
2.2 Kinect のデータを取得する流れ.....	4
2.2.1 Depth センサー.....	4
2.2.2 BodyIndex .....	5
2.2.3 人物姿勢 .....	5
2.2.4 手の状態の認識.....	6
3.方法.....	6
4.今後の展望 .....	10
5.謝辞.....	10
6.参考文献.....	10

## 1.はじめに

2010年に米国マイクロソフトから発売された Kinect v1 は、Xbox 360 用の入力システムとして登場した Natural User Interface (NUI)、つまり人とコンピュータ間対話の画期的な新技術として、世界中の開発者や研究者から注目された。現在では、2014 年販売された Kinect v2 はさらに進化して、指先と親指までの情報も正確に取ることができた。

また、最近では VR メガネが増えた。VR メガネが幻想の景色にタイプしたと感じた、でもコントローラを握ると又現実に戻ると感じる。

そこでタイプした感を強くするために、本研究では Kinect を用いて手による入力デバイスに着目して研究を行う。

## 2.Kinect の概要

Kinect (キネクト) はマイクロソフトから発売されたジェスチャー・音声認識によって操作ができるデバイス[1]。動力学を意味する "Kinetics" と、繋がりを表す "Connection" を組み合わせた造語である。キャッチコピーは「カラダまるごとコントローラー」、開発コード名は「Project Natal」、「NUI」(ナチュラユーザーインターフェイス) の一つである。

現在では Kinect v1 と Kinect v2 の二つがある。

### 2.1 Kinect v1 と Kinect v2 の区別

#### 2.1.1 Kinect の仕組みと原理



図 1.kinect v1

Kinect v1 の Depth センサーは、投光した赤外線パターンを読み取り、パターンのゆがみから Depth 情報を得る「Light Coding」という方式が採用されている。そのため、Depth センサーは赤外線パターンを投光する IR プロジェクター (左) とそれを読み取る IR カメラ (右) に分かれている。また、Depth センサーの間には Color カメラが搭載されている。

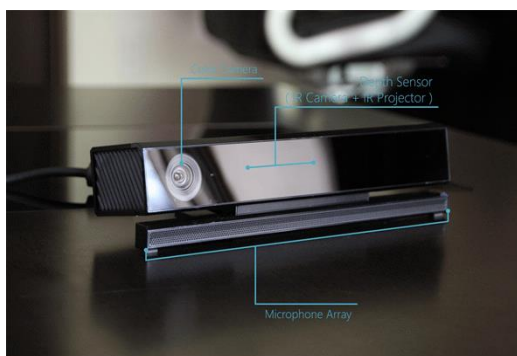


図 2.kinect v2

Kinect v2 の Depth センサーは、投光した赤外線が反射して戻ってくる時間から Depth 情報を得る「Time of Flight (TOF)」\*2 という方式が採用されている。Depth センサーは、外観からは見えないが、Color カメラの隣に赤外線カメラ (左) とパルス変調された赤外線を投光するプロジェクター (右) が搭載されている。

### 2.1.2 Kinect の仕様

Kinect v1 と Kinect v2 のセンサーの仕様比較を表 1 に示す。

表 1 Kinect v1 と Kinect v2 の特徴比較

特徴	Kinect v1	Kinect v2
Color 画像	640×480 @ 30fps	1920×1080 @30 fps
Depth 画像	320×240	512×424
最大奥行距離	~4.5M	~4.5M
最少奥行距離	40cm (Near Mode)	50cm
水平視角	57 度	70 度
垂直視角	43 度	60 度
チルトモーター	あり	無し
骨格定義数	20 関節/1 人	25 関節/1 人
最大追跡対象 スケルトン	2 人	6 人
手指の検出	なし	親指と指先
手のポーズ検出	グーとパーのみ	グー、チョキ、 パー
音声入力	あり	あり

### 2.1.3 Kinect の開発環境

Kinect v1 と Kinect v2 の開発環境の比較表を表 2 に示す。

表 2 Kinect v1 と Kinect v2 の開発環境比較

	Kinect v1	Kinect v2
OS	Windows7 以降	Windows 8 以降 (x64)
コンパイラ (Compiler)	Visual Studio 2010 以降	Visual Studio 2012 以降
接続端子 (Connector)	USB 2.0	USB 3.0
CPU	Dual-Core 2.66GHz	Intel Core i7 3.1GHz 以上
GPU	DirectX 9.0c	DirectX 11.0
RAM	2.0GB 以上	4.0 GB 以上

Kinect v2 のソフトウェア開発キットとして Kinect for Windows SDK v2.0 が無償で提供されている。Kinect v1 の Kinect for Windows SDK v1.x とは互換性はない。

## 2.2 Kinect のデータを取得する流れ

Sensor から Source を取得、Source から Reader を開き、Reader から Frame を取得、Frame からデータを取得するという流れ。

Kinect v1 と比べれば、Reader というステップが追加された。Reader は 1 つの Source に対して複数開くことができる。この仕様により、マルチスレッドのアプリケーションでは取得したデータを別のスレッドにコピーするといった処理が不要になる、複数のアプリケーションでは同じセンサーからデータを取得できるなどの利点がある。

### 2.2.1 Depth センサー

Kinect には Depth センサーが搭載されており、センサー面からの距離情報、言わば Depth データという情報を取得できる。

Kinect v1 では、投光した赤外線パターンを読み取り、パターンのゆがみから Depth 情報を得る、言わば Light Coding という方式の Depth センサーが搭載されていた。

Kinect v2 では、投光した赤外線パルスが反射して戻ってくるまでの時間から Depth データを得る、言わば Time of Flight(ToF)という方式の Depth センサーに変更されている。

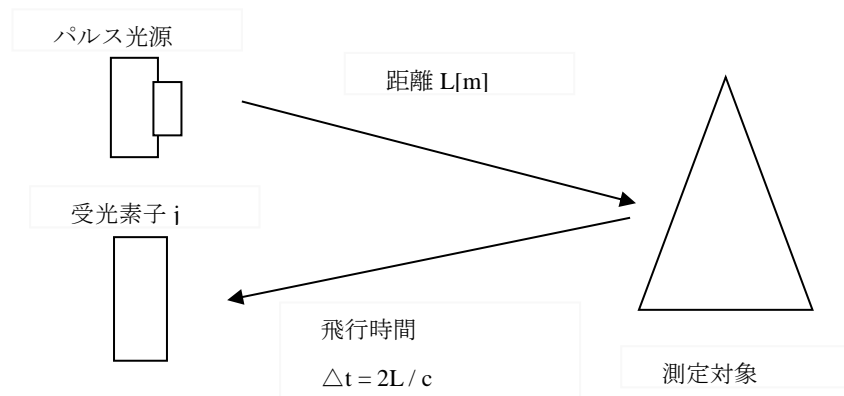


図 3.Time of Flight(ToF)



図 4.Depth 画像

### 2.2.2 BodyIndex

BodyIndex とは人物領域である。

Kinect は取得した Depth データを基にした人物領域を取得できる。

人物領域は Depth データを基にしているため、Depth センサーの解像度と同じになる。Kinect v2 では Depth センサーの解像度や分解能が大幅に向上しているため、Kinect v1 に比べ、手指などの細かい部分も人物領域として崩れることなく取得できるようになった。

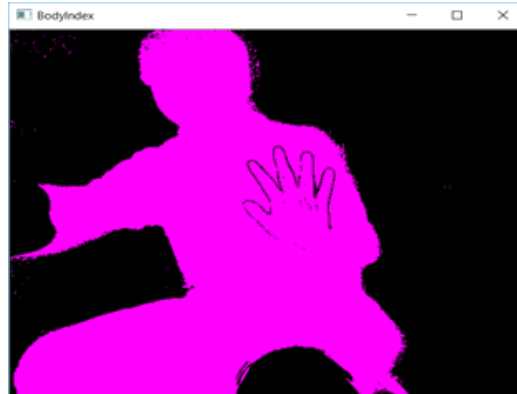


図 5.BodyIndex 画像

### 2.2.3 人物姿勢

Kinect は Depth (=センサー面からの距離情報) や BodyIndex (=人物領域) を取得し、これらのデータを基にした人物姿勢を取得できる。

Kinect の人物姿勢は、膨大な数の姿勢情報を基に学習された識別器に人物領域の距離情報を入力することで推定している。

背景となる技術は難しく感じるかもしれないが、開発者は Kinect SDK を通じて簡単に人物姿勢を取得し利用できる。

人物姿勢データは、頭、手、足などの 3 次元位置を得ることができ、それを基にジェスチャー認識などを実装できる。Kinect v2 では Depth センサーの解像度や分解能が大幅に向上しているため、Kinect v1 に比べ、指先や親指なども取得できるようになった。Kinect v2 では、これらの詳細なデータを検出可能な 6 人全てで取得できる。

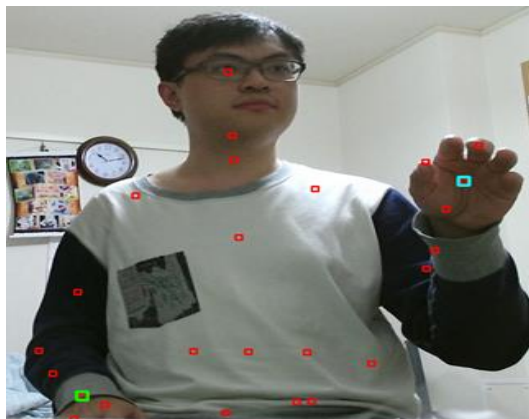


図 6. 人物姿勢

#### 2.2.4 手の状態の認識

Kinect v1 の手の状態はパーとグーの 2 種類であった。

Kinect v2 は指先や親指などを取得できるので、チョキの状態も取得できるようになった。

手の状態を色で表示する。

不明の時は青、追跡不可の時は緑、パーの時はシアン、グーの時はマゼンタ、チョキの時は黄色と設定する。

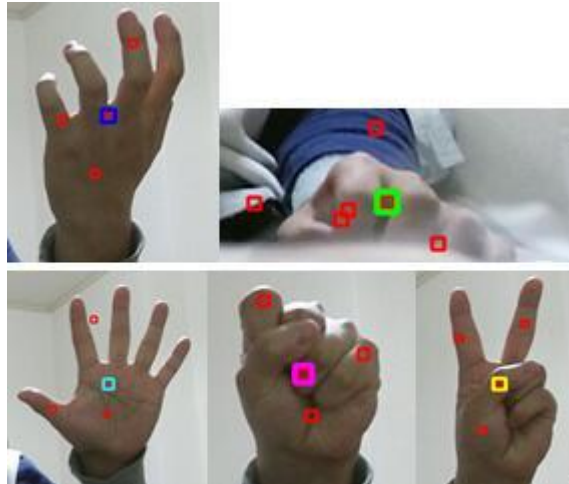


図 7. 手の状態

### 3.方法

球体を掴みながら移動する、放下したい場所に放置するプログラムを作る。

具体的に図 8、9、10 のフローチャート通り作る。

球体のコントロールは図 11、12、13、14 に表する。

グーの時は球体を移動する。

パーの時は球体を放置する。

チョキの時は球体を固定する。

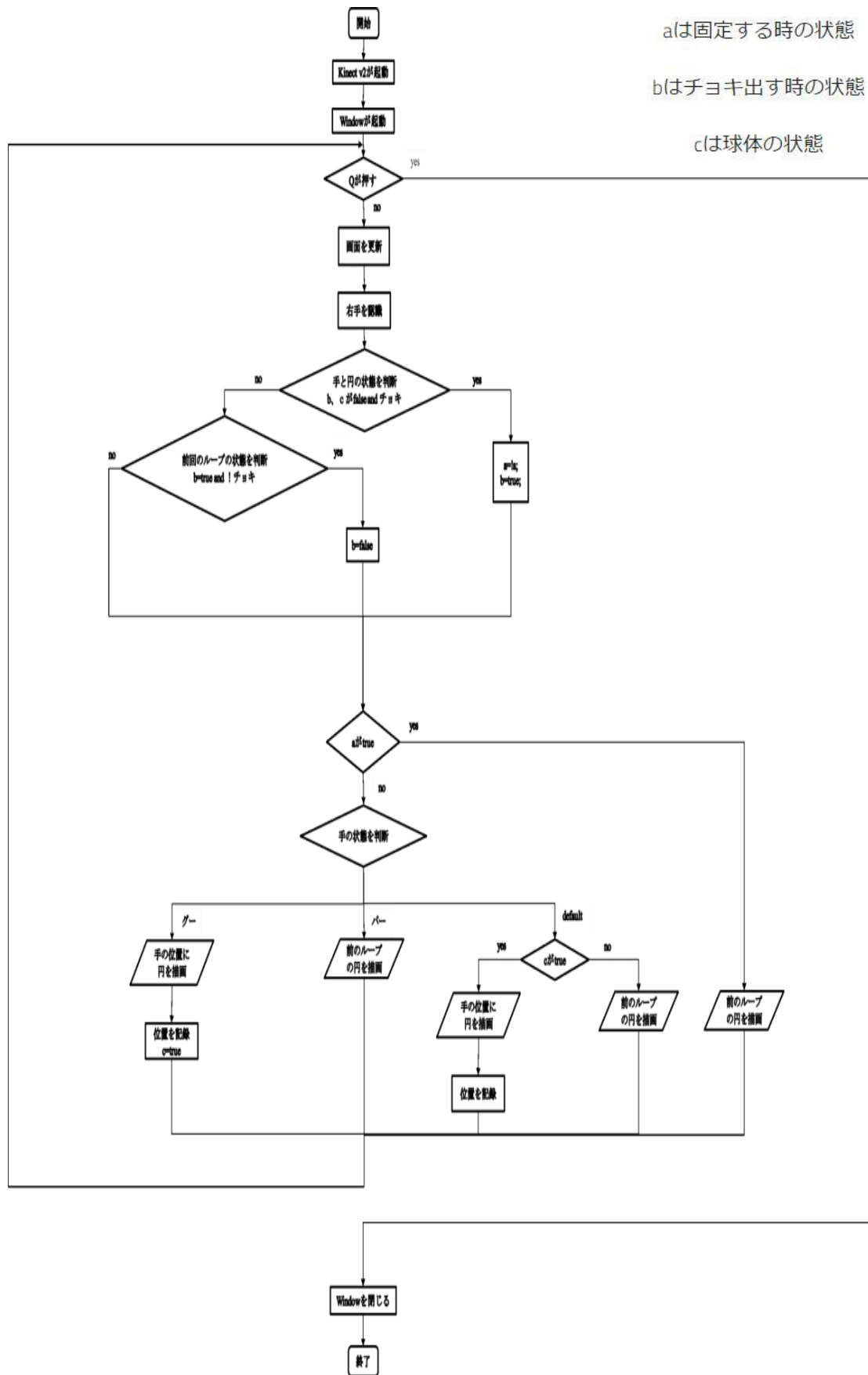


図 8.フローチャート



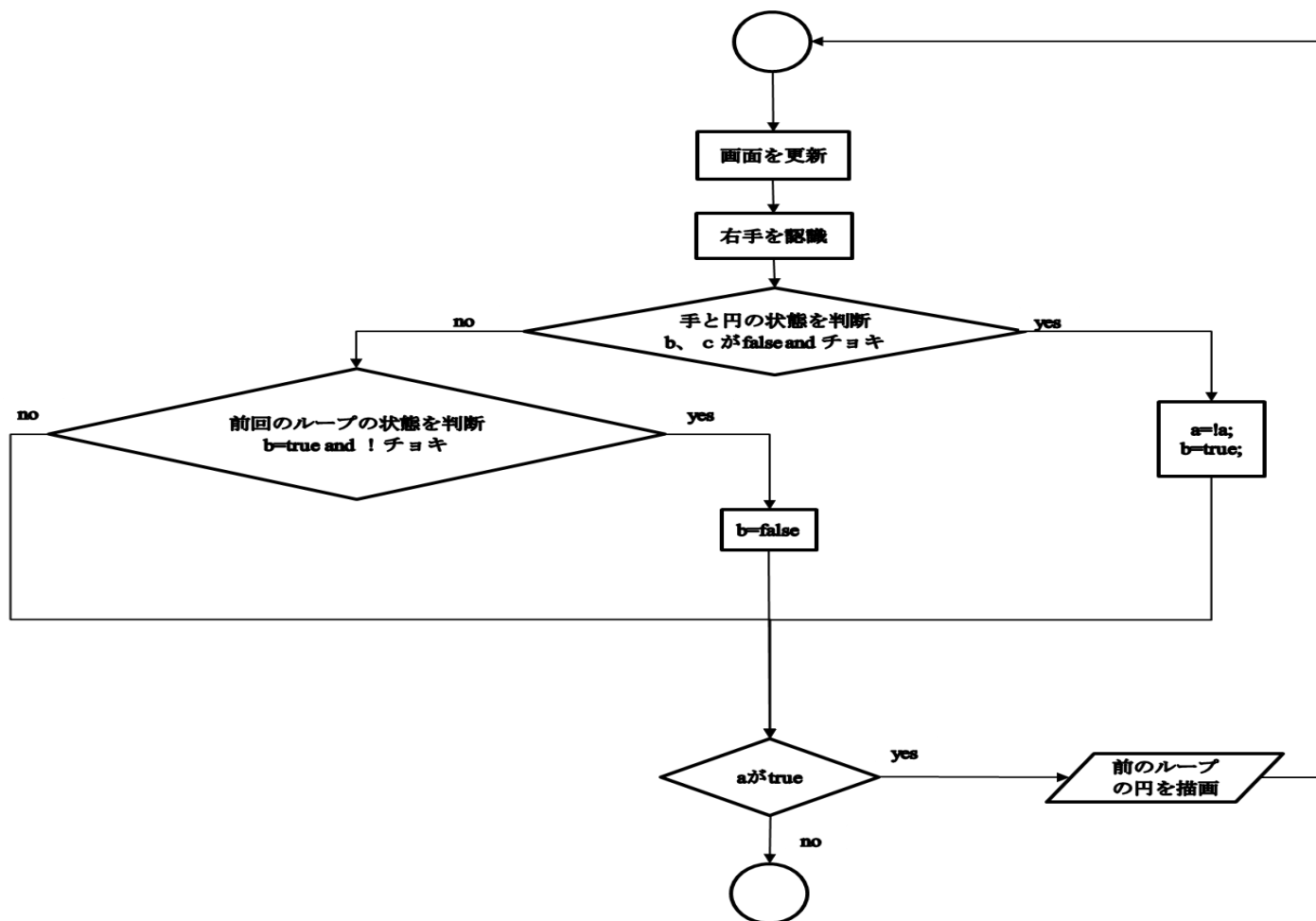


図 9.フローチャートの上の部分

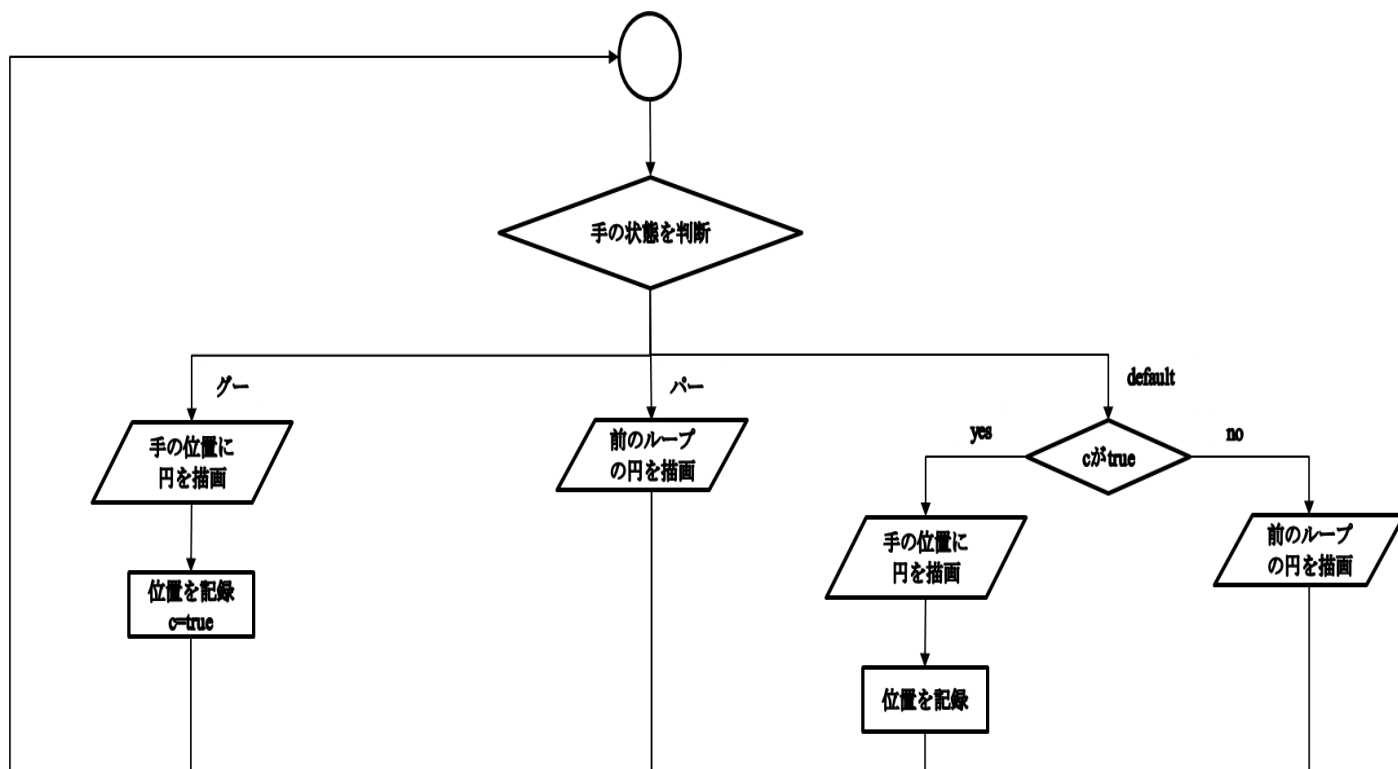


図 10.フローチャートの下の部分



図 11.球体を平面で移動する

グーの時は球体を移動する。



図 12.球体を前後で移動する

グーの時は球体を移動する。



図 13. 球体を放置する

パーで球体を放置する、でもグーの時は物体を固定できない。



図 14.球体を固定する

1 回目のチョキで固定する、グーの時も移動できない。2 回目のチョキで固定を解除する。

## 4.今後の展望

チェス、囲碁、将棋など、リアルと同様に手で駒を操作する、マウスを手で操作する、PPTを手で操作するなどできます、マルチタッチするなどいろいろなことができる。

キネクト v2 は手だけではなく、表情データも取得することができる。表情データを用いてネットゲーム中相手の表情も表現できる、表情による入力デバイス、AI の知能を更に高めるなどいろいろなことができる。

## 5.謝辞

丁寧にご指導していただき、また機材を買ってくださった高橋教授に心より感謝の意を表します。また日頃より貴重なご意見、ご指導をしてくださった中村高橋研究室の先輩方に深く感謝いたします。

## 6.参考文献

- [1] 葉師寺 国安, “Kinect v2 プログラミング入門”, 株式会社インプレス, 2014
- [2] 藤本 雄一郎, 青砥 隆仁, 浦西 友樹, 大倉 史生, 小枝 正直, 中島 悠太, 山本 豪志朗 “OpenCV3 プログラミングブック”, 株式会社マイナビ, 2015
- [3] 新田 善久, “NtKinect: Kinect V2 C++ Programming with OpenCV on Windows10”, <http://nw.tsuda.ac.jp/lec/kinect2/>, 2016