

## キネクトを用いた障がい者支援用入力デバイスの試作

宮崎英一，坂井聡\*，谷口公彦\*\*，野田知良\*\*\*，  
大野香織\*\*\*\*，篠原智代\*\*\*\*  
(技術教育)(特別支援教育)\*，(香川県立高松養護学校教諭)\*\*，  
(香川県立聾学校教諭)\*\*\*，(かがわ総合リハビリテーションセンター)\*\*\*\*

760-8522 高松市幸町 1 - 1 香川大学教育学部  
\*760-8522 高松市幸町 1 - 1 香川大学教育学部  
\*\*761-8057 高松市田村町1098 香川県立高松養護学校  
\*\*\*761-8074 高松市太田上町513- 1 香川県立聾学校  
\*\*\*\*761-8057 高松市田村町1114番地 かがわ総合リハビリテーションセンター

## The trial production of the disabled person Input device for Kinect.

Eiichi MIYAZAKI, Satoshi SAKAI, Kimihiko TANIGUCHI, Tomoyoshi NODA,  
Kaori OONO and Tomoyo SHINOHARA

*Faculty of Education, Kagawa University, 1-1 Saiwai-cho, Takamatsu 760-8522*

*\*Faculty of Education, Kagawa University, 1-1 Saiwai-cho, Takamatsu 760-8522*

*\*\*Kagawa Prefectural Kagawa Special education school, 1098 Tamura-cho, Takamatsu 761-8057*

*\*\*\*Kagawa Prefectural School for the Deaf, 513-1 Otakami-machi, Takamatsu 761-8074*

*\*\*\*\*Kagawa General Rehabiritation Center, 1114 Tamura-cho, Takamatsu 769-2302*

**要旨** 本研究ではゲーム機の入力デバイスとして開発されたキネクトを応用する事で，マウスやキーボードといった従来の一般的な入力デバイスが利用困難な方でも使用可能な入力支援システムを試作した。このシステムのソフトウェア部分を自作する事で，ユーザの入力を音声認識や顔の動作といった人間が持つ自然的な操作で制御が行えるようになった。またマウス等の代替だけでなく，赤外線リモコンと組み合わせる事で，日常生活に用いられる家電製品の制

御が行えるようになった。このためゲーム機の入力部分を応用する事で，従来の同様の機能を持つ高価な機器が，安価で入手しやすい形で提供できる事が示された。

キーワード 障がい者支援，キネクト，音声認識，モーション，NUI

## 1. はじめに

キネクトは2010年にマイクロソフトから，コンシューマゲーム機（Xbox360）の体感的なユーザインタフェースの普及を目指し，新しく開発された入力デバイスである。これは従来のようなマウスやキーボード等のインタフェースを介さずに，音声入力や，ユーザの動きから入力を行うモーション・センシングを利用し，ゲーム等の操作を行うものである。このユーザインタフェースはNUI（Natural User Interface）と呼ばれ，従来のマウス等を用いたGUI（Graphical User Interface）に置き換わるものとして普及が期待されている。

現状では，入力方法の問題でNUIの操作方法では，キーボードから文章を入力するような操作等には不向きである。しかし，ユーザの動作と制御内容を直感的に関連づけて行う事ができるので，簡単な操作に関しては，操作方法の学習が不必要なだけでなく，何らかの原因でマウスやキーボードの操作が行えないユーザにとっても操作上，大きなメリットをもたらすと考えられる。

本研究では，これらのユーザインタフェースの特性を生かし，従来のマウスやキーボードといったGUI環境下で操作が行えない人でも，コンピュータの入力や，家電製品の制御が行える入力支援システムの試作を目指すものである。

## 2. キネクトについて

### 2.1 開発環境

当初，キネクトはゲーム機に附属した入力デバイスとして市販されていたが，NUIの持つ潜在的有効性と一般家庭への普及の素早さから，独自にコンピュータと接続して新しい入力インタフェースとして使用するという形が広まってきた。従来のモーションキャプチャーの機器<sup>1</sup>は殆どがプロ用という事もあって非常に高価（100万円～）であり，気軽に使用できるというものではなかった。しかしキネクトは元来がゲーム機の入力デバイスとして開発されたため，安価（1万4千円程度）であり，誰でもが購入可能な価格帯まで下がってきた事もその原因の一つであろう。これらの結果から，今後もより広く一般家庭レベルにおいてゲーム機以外の使用も十分実用的であると予想される。

更に，公式な開発環境として，Kinect for Windows SDK<sup>2</sup>はバージョン1.8までが提供されており，バージョンアップに伴う新しい機能が次々と追加されている。最新のバージョン<sup>3</sup>ではリアルタイムで正確な3次元レンダリングを行う“Kinect Fusion”も，従来よりその機能を大幅に強化し，キネクトだけで実空間の実物から3Dモデルの取り込みが可能になるので，3Dプリンタの普及をも後押ししている。

### 2.2 障がい者への対応

キネクトを用いた障がい者入力支援システ

ムとしては、先行研究として東京大学先端科学技術研究センターと日本マイクロソフトが既に開発を進めており、幾つかの報告<sup>4, 5</sup>, や、既にこれを用いた市販製品<sup>6</sup>も発表されている。これらのシステムは学校や家庭内といった現場での使用を目的としているので、キネクトだけでなく、これから外部機器を制御する赤外線装置やUSBインタフェースまでも含めた形で市販されている。このため多くの場合、使用者はこれらのパッケージを購入するだけで、入力手段だけでなく、アウトプットとして様々な機器の使用が可能になる環境まで提供されている。

しかし反面、これらのシステムは全体がパッケージとして市販されていることから、ソースコードの入手が困難なため、こちらで既に使用している機器に合わせて制御機器を追加する事は困難であるという問題点があった。

### 2.3 使用環境に基づく開発ターゲット

本研究では先行研究に習い、キネクトを用いたNUIを提供する事で、障がいを持った方の入力手段を補助し、赤外線リモコン等で制御される家電製品の制御を行うものである。これはキネクトのシステムのコアとなる顔認識や音声認識の部分をソースコードから自作する事で、赤外線リモコン等手持ちの障がい者支援入力デバイス<sup>7</sup>の制御を行う事で達成できると考えた。

本研究では赤外線リモコンを用いた障がい者支援の入力デバイスの操作対象を、多目的な展開が可能と予想されるビデオの再生とした。これは教育現場でも単なる学習用教材として用いられるだけでなく、スイッチ機能を応用した学習といった基本的な部分にも使用されている。その他、家庭内においても余暇の分野で多くの利用が期待でき、本人の操作

でビデオの操作ができれば、本人だけでなく、サポートをする側の負担軽減も期待できるからである。

元来、キネクトはゲーム機器のユーザインタフェースとして使用されていたため、モーションキャプチャーとして全身の骨格（ボーンモデル）を撮影し、この骨格情報からユーザの動作を解析する<sup>8</sup>ものである。しかし、障がいをもった方の利用を想定した場合、全身を使用した入力は運動機能等の制限から、運用上からは全身骨格モデルの適合は困難な場合が多いと考えられる。

そこで本研究では関節の骨格情報を利用しない、幾つかの入力環境を提案する。この場合、キネクトのカメラセンサーやマイクロフォンアレイ機能を利用した1)顔追跡、2)音声認識、3)手指操作等の入力方法が考えられる。

1)はカメラで撮影された使用者の顔画像を解析し、顔の傾きや目の瞬き、口の開閉、眉毛の上げ下げ等をユーザの制御入力情報とするものである。この時、画像処理として単なる差分方法を用いずに顔認識を行えば、口の開閉を入力とした場合、使用者が利用中に顔を動かした場合（例、回転・移動運動等が発生した場合）でも口が顔のパーツとして認識されているので、移動に伴う誤差が発生しないという利点がある。

2)は古くからある入力方法で、マイクで録音した音声をリアルタイムに解析し、解析された語彙からユーザの入力情報を判別するものである。これも使用者の予期しない体の動きに影響を受けないので、発声・発話が安定していれば、使いやすい入力方法の提供が可能である。ただし、音声認識エンジンが日本語に対応している必要があり、従来はキネクトを使用した音声認識を行う上で、これが

問題になっていた。

3) は撮影された画像から手の移動や、指の上げ下げ、回転等の動作を判定し、これを入力情報とするものである。この時、測定対象の空間位置を測定対象とすると、測定中は手の位置を空間的に固定する必要がある、運動機能に制限のある方の場合、操作に支障が出てくる可能性も十分考えられる。

初期の開発環境ではSDKのサポートから1), 2) の開発は困難であったが、Kinect for Windows SDKのバージョン1.5からは顔追跡と日本語をサポートした音声認識が正式にサポートされるようになり、ユーザ側のプログラムからこれらのセンサー値を直接読み込めるようになった。3) の操作に関しては、キネクトの認識範囲は40センチ～(Xbox用Kinectは80センチ～) というのもあってか、現在の最新の(Kinect for Windows SDK バージョン1.8) でもサポートされておらず、現在でもユーザプログラムからのセンサー利用が困難な状況である。よって本研究ではキネクトのユーザの入力方法として、顔追跡(顔の回転、傾き)を利用した入力と日本語音声入力をを用いた入力方法の2つの方法を採用する。

本研究では上記で述べた理由からキネクトを用いた家電製品の制御対象を赤外線リモコンを用いたビデオとしている。ビデオの制御にはいろいろな赤外線制御コマンド(命令)があるが、利用者の利用状況を考慮して、ビデオの再生を目的とした。この場合、利用者の要求度が高いコマンドを制御対象とし、再生・(停止)、早送り(進む)、スロー再生(戻る)、インデックス(一覧表示)の4つのコマンドを音声認識および顔認識を用いた入力支援システムでの制御コマンドとした。

### 3. 音声認識

#### 3.1 音声認識の環境準備

キネクトで日本語音声認識を行う場合、Kinect for Windows SDKをインストールする以外に、別途、Microsoft Speech Platform Software Development Kit (Version 11) と、Kinect for Windows Language Packs v1.5.0の日本語用ロケールが必要になるので、予めインストールしておく必要がある。

今回は日本語対応とする事と、新たに音声コマンドを追加するためにサンプルコードとして記載されている「Speech Basic-WPF」と文献<sup>9</sup>の2つを参考とし、これに赤外線リモコンを制御するプログラムを組み込んだものを試作した。キネクトで日本語認識を行う場合、ソースコード内に日本語の音声認識用の単語を追加するだけで簡単に使用できるという特徴がある。

表1に本研究で試作した赤外線リモコンの制御に用いるために登録した日本語音声コマンドを示す。ここでのキーワードは赤外線リモコンの制御命令に関連付けられており、このキーワードに対して複数の音声コマンドを登録する事ができる。ここでは「ビデオを見る」という動作に対して「さいせい」・「みる」と2つの音声コマンドを登録している。この方法で、複数の音声コマンドが登録できるので、一般的な単語を複数登録しておけば、実際の音声コマンドを知らなくても、ユーザが直感的に思いつく単語で機器制御が可能にな

表1 登録された音声コマンド

キーワード	日本語音声コマンド
FORWARD	インデックス, いちらん
BACKWARD	さいせい, みる
LEFT	おくり, おくる
RIGHT	もどし, もどる



図1 音声認識プログラム画面

り、ユーザに対してより使いやすい環境を提供できると考えられる。しかし、現状では「単語」の認識だけが行っておらず、「テレビのチャンネルをNHKに変える」というような文章から自動的に単語を切り出して認識するような形式にはなっていない。しかし今後は、よりユーザの利便性を向上させ、使いやすいインタフェースを提供するために、この部分の改良も必要となる可能性がある。

図1にキネクトを用いた音声認識プログラムの実行画面を示している。同図に示された「再生」という文字は、音声認識により認識された「音声コマンド」を示している。例えば、このまま続けて「停止」と発声すると、画面の文字は自動的に「停止」と表示され、赤外線リモコンの停止命令が実行される。

このようにプログラムを実行すると、キネクトのマイクから入力される音声を自動的に認識し、その音声に応じた命令で赤外線リモコンの制御までがシームレスに行われている。このため、一度プログラムが実行されてしまえば、ユーザは全くコンピュータに触る事なしに連続して音声認識が行えるので、ユーザに負担をかけないシステムとなっている。

また音声認識の精度に関しても、通常の会話程度の発声ならば殆ど誤認識が発生せず、快適に制御を行う事ができた。更に音声認識

中に他の人が横から音声を発しても、これも認識され、キネクト1台で複数のユーザからの音声認識を行う事も可能であった。このため学校等の環境下で複数のユーザが使用する場合でも、問題なく使用できる事もわかった。

### 3.2 赤外線リモコンの制御

本研究では、キネクトで認識された制御信号を基に、赤外線リモコンの発光制御を行う。今回は、図2に示すように赤外線リモコンのハードウェアをビット・トレード・ワンの「USB接続 赤外線リモコン」を用いた。これは、学習タイプの赤外線リモコンであるが、USBで接続されたコンピュータからプログラムで制御が行える。さらに、赤外線送信ライブラリが提供されているので、自作プログラムからでも簡単に制御を行う事が出来るという利点がある。本研究では、キネクトを制御している自作のC#のプログラムから他のプロセスを起動する「System.Diagnostics.Process.Start」を用いて赤外線制御プログラムを別プロセスとして呼び出している。ところが、実際にテストをしてみると、入力トリガーが連続して赤外線信号を送信してしまい、リモコンのボタンを連打したような状態になってしまった。

そこで回避策として赤外線信号を送信後、

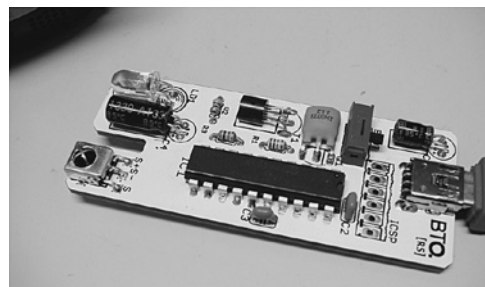


図2 赤外線リモコン

「System.Threading.Thread.Sleep」を発行して、親スレッドを強制的に待機状態とした。この待機時間は実際の環境下で幾つかの時間パラメータをテストしてみて0.5秒とした。このため、実際の運用ではコマンドが実行されてから0.5秒の待機時間があるが、体感上支障のない応答性となっている。

また赤外線リモコンの制御には別プロセスを起動する時に、「Arguments」を利用して、コマンドラインから別プロセスに赤外線リモコンの送信データを渡している。このため、制御対象となる赤外線リモコンの命令（コマンド）を変更する場合、親プロセスにハードコートされている赤外線データ部分を変更するだけで済む。その結果、赤外線リモコンを送信するプログラムに手を加える必要がないので、操作命令を修正する場合でもプログラム本体の変更だけで柔軟に対応させる事が可能となった。ここでは赤外線信号の送信プログラムはWindowsコンソールアプリケーション<sup>10</sup>のソースコードをそのまま利用させて頂いている。

### 3.3 顔認識

キネクトを用いて顔認識を場合、Kinect For Windows SDKを使用するのが一般的である。同SDKには、Microsoft Face Tracking Software Development Kit for Kinect for Windows が同梱されており、これを利用する事で、ユーザの作成したアプリケーションからでも顔認識を行う事が可能になる。Face Tracking SDKは、“for Kinect for Windows”という名前の通り、Kinectセンサーと共に利用するように設計されており、深度情報や骨格情報も利用した高速・高精度の認識が可能である。具体的には表2に示した内容が実現可能となる。

本研究では、同表の4)の機能を利用し、顔の向きを入力トリガーとした。ここでは

表2 追加された機能

1) 画像情報からの顔の検出
2) 顔の追跡
3) 顔の部位各点の座標の取得
4) 顔の向きの取得
5) 表情（口を開いている、眉の上げ下げなど）の取得
6) 顔の3Dモデルの取得

実際の環境で他の機能もテストしてみたが、5)等ではあまり認識率が上がらず、使用者の操作負担の軽減や再現のし易さ等を考慮して4)を入力トリガーとしている。本研究で規定した顔の傾きと赤外線リモコンから発行される制御コマンドを図3に示す。ここでは、一番操作に負担が少ないと思われる顔の左右振り（左右を向く：ピッチ角測定）という動きに、使用頻度が高い再生・停止と再生ビデオを選択する一覧（インデックス）を割り当てた。また早送りと巻き戻しには顔の上下の振り（頷く・仰ぐ：ヨー角測定）を割り当てている。実際にテストをしてみると、上下に振った場合は、顔認識がロストする場合が多く、余り大きな動作でない方が良い事が分かった。またキネクトに接近しすぎても、

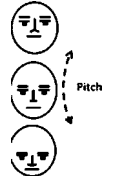
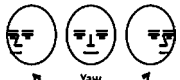
顔の傾き		制御コマンド	
X軸 Pitch angle		頷く	進む
		天を仰ぐ	戻る
Y軸 Yaw angle		右を向く	再生（停止）
		左を向く	インデックス

図3 顔の認識<sup>11</sup>とコマンド対応

## キネクトを用いた障がい者支援用入力デバイスの試作



図4 顔認識プログラム

上半身の情報が欠落してしまうので、顔認識そのものが行えなかった。

試作したプログラム内では、FaceTrackFrameオブジェクトから顔の傾きに関する情報(Rotation)を抽出し、顔面モデルのX軸、Y軸に対する回転角度を測定している。ここでは顔の傾きにに応じて $-90\text{度} \leq \text{Rotation} \leq 90\text{度}$ の範囲で変化するので、これを用いて「顔が傾く⇒Rotation値検出⇒制御コマンド発行」の実行が可能になった。赤外線信号の制御に関しては上記の音声認識で説明したプログラムをそのまま応用している。

図4に実際に顔認識を行っているプログラムの画面を示している。ここでは、キネクトにより認識された顔が赤い線で描画されたポリゴンモデルとして表示されている。このモデルでは、顔の輪郭だけでなく、目・口・鼻といった各部分のパーツを認識しており、これらの位置情報を用いてポリゴンモデルを表示している。同図の上部にはキネクトで測定されたX軸(ピッチ角)、Y軸(ヨー角)がリアルタイムで表示されている。この値が規定値を越して変化すると、赤外線リモコンから制御信号が発光され、機器の制御が行われる。

図5には試作したプログラムを用いてキネ

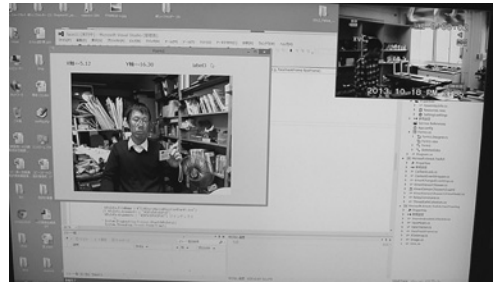


図5 b) 顔認識による制御画面



図5 a) 赤外線リモコンによるカメラ制御

クトで実際にビデオの再生制御を行っている様子を示している。同図a)には赤外線リモコンとこれにより制御されるビデオカメラ(ビデオ再生用)を示した。ここでは写真撮影の都合上、両者をきわめて近い位置に配置しているが、実際の運用では、3m程度離しても、制御が行えた。

また同図b)には顔認識のプログラム画面とこれにより制御されているビデオ再生画面を示している。本来の使用方法では、再生されるビデオ画像はTVなどの機器に接続し、パソコンの画面に表示する必要はないが、稼働テストの都合上、顔認識とビデオ再生画面を同時に表示している。このシステムを用いて認識精度をテストしてみたが、顔の傾け方に慣れてしまえば、実用上は十分な精度がある事が分かった。

#### 4. おわりに

研究ではゲーム機の入力デバイスとして市販されているキネクトを応用する事で，マウスやキーボードといった従来のインタフェースでは操作方法の問題で利用出来なかった方でも，これを補完する形のインタフェースとして使用可能な事が示された。具体的には音声認識や画像処理を入力に応用する事で，顔の傾きといった従来では困難だったインタフェースが簡単に実現可能となった。

さらにこのプログラムを自作する事で，家電製品の赤外線リモコンの制御が行えるようになったので，ビデオの制御が可能となり，日常生活を補助する環境を簡単に構築できるようになった。今後は，実際に学校や家庭で使用してもらい，その問題点の改良等に努めたいと思う。

#### 5. 謝辞

本研究は，平成25年度科学研究費補助金（基盤研究（C））「運動機能及び発達障害をサポートする生活・学習支援ワンデバイスシステムに関する研究」（課題番号24500648）の一部として行われたことを記して謝意を示す。

#### 6. 参考文献

<sup>1</sup> <http://www.mocap.jp/> 標準システムで126万円～

<sup>2</sup> <http://www.microsoft.com/en-us/kinectforwindows/>

<sup>3</sup> <http://www.microsoft.com/en-us/kinectforwindowsdev/Downloads.aspx>

<sup>4</sup> Kinectを活用した障害者向け活動支援ソリューション「OAK」，東大先端研と日本MSが共同開発 <http://monoist.atmarkit.co.jp/mn/articles/1210/03/news070.html>

<sup>5</sup> 東大先端研と日本マイクロソフト，Kinectで障害者の意思表示を支援するソリューション <http://japan.internet.com/webtech/20121004/1.htm>

<sup>6</sup> どこでもスイッチ【OAK】—Observation and Access with Kinect— <http://www.assist-i.net/modules/products/oak.html>

<sup>7</sup> マイクロコントローラを用いた障がい者用 入力補助デバイスの試作（The trial production of the disabled person input auxiliary device for microcontrollers），宮崎英一，その他，香川大学教育学部研究報告第Ⅱ部，63，1，pp.13-20，2013

<sup>8</sup> Making Things See—KinectとProcessingではじめる3Dプログラミング，Greg Borenstein，オライリージャパン，2013

<sup>9</sup> Kinect for Windowsアプリを作ろう，日向俊二，カットシステム，2012

<sup>10</sup> [DSAS開発者の部屋]，<http://dsas.blog.klab.org/archives/52097996.html>

<sup>11</sup> <http://msdn.microsoft.com/en-us/library/jj130970.aspx>