

# 動作・音声認識によるプレゼンテーション用PC操作支援システム A Hands-free Input Interface of Presenter Manipulating PC for Presentation

12N8100010J 初 蓄  
情報工学専攻 牧野研究室  
2014年3月

## あらまし

プロジェクタを使うプレゼンテーション中に、発表者の両手を解放し、プレゼンテーション室内で自由に移動する環境を作ること、本研究の目的とする。そのために、本研究ではプレゼンテーションの時、ユーザの両手を解放し、室内で自由に移動する環境を実現するシステムを構築する。

キーワード: 入力インタフェース, ハンズフリー, Kinect, 教育支援

## 1 序論

現在、PCと投影技術の開発より、大画面でグラフィカル資料を用いた効果的なプレゼンテーションができる。しかし、利用者はPCから離れることができず、自分の移動範囲が制限されることである。もし、離れた場合には、PCの遠隔操作ができる装置が必要になる。現在、PC画面をプロジェクタ等で投影して受講者に見せながらプレゼンテーションを進行する講演者を支援するためのツールが出ている。この種の機器で代表的なのはプレゼンポインタ、タブレット端末とワイヤレスマウスである。プレゼンポインタには、レーザー光を射出する一般的なものから、光を発する代わりに画面上に描画するものまで、様々な商品が存在する。しかし、操作機能としては、マウスの左右ボタンのみに対応するものがほとんどであり、他のアプリケーションへの切り替えなどはできない。これに対して、タブレット端末はほとんどの操作を画面に指を触れて行うことができる。[2]。しかし、電力消費量と重量が理由で長時間の連続使用は難しい。一方、ワイヤレスマウスでは複雑な操作が簡単だが、でも、操作方法のために、自由に移動することができない。本研究では、プロジェクタを使う授業またはプレゼンテーションの時、ユーザの両手を解放し、室内で自由に移動する環境を作るために、新しいプレゼンテーション支援システムを提案する。

## 2 提案システム

### 2.1 概要

そのために、図1に示す様に、従来の入力手段(マウス、キーボード、Air Beamなど)に代わって、Kinectを入力手段とする。提案システムでは、Kinectを通じてユーザの動作・音声を取得・認識して、登録済みの動作・音声の場合には対応する操作指示をWindows OSを通じてアプリケーションに入力する[1]。対象とするアプリケーションとしては、プレゼンテーションでよく使われているパワーポイント、WordとAdobe Acrobat Readerを選定する。

### 2.2 システムの流れ

システムの流れを図2に示す

**Step1:** システムを起動する

**Step2:** ユーザ認識とロックする

画面に映る人物が、ユーザとしての登録動作をするまで、次に遷移しない。

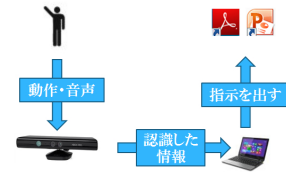


図 1: システムの全体像

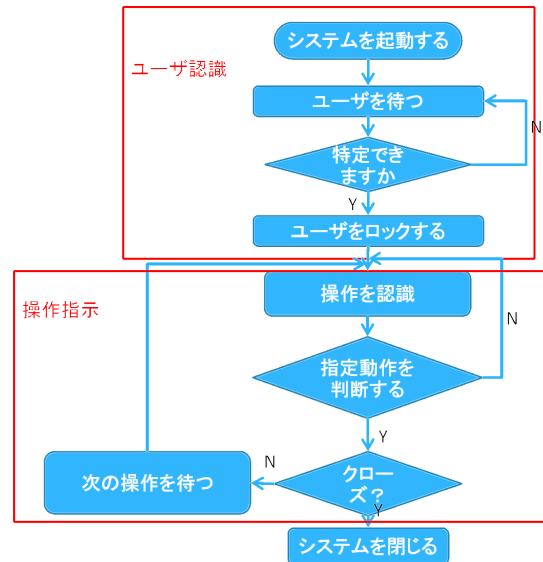


図 2: システムの流れ

### Step3: 動作認識

音声の認識とそれに伴うコマンド決定が抜けている。動作とマッチングしているアプリケーションコマンドを実行する。

### Step4: システムを閉じる

所定の動作により、システムを終了する。

### 2.3 ユーザの判明

提案システムは、授業・会議等での発表で用いることを想定している。この場合、多数の人が同じ部屋にいるため、ユーザ以外の人物の動作や音声による誤作動の恐れがある。このため、提案システムでは初めにユーザを特定し、特定された人物以外の動作や音声を認識しない。すなわち、特定された人物のみが提案システムを通じてアプリケーションを操作できる。この権限を渡す前に、システムは他の動作を認識しない。プレゼンテーションする時、講演者は歩いてながら講演していることを含めて考えると、ユーザを判明する条件として、距離だけでは足りない。そこで、システム

は、左手を前に伸ばす人をユーザとして認識し、ユーザをロックする。その後、提案システムはロックしたユーザの動作および音声のみを認識する。この結果、他人がその後同一動作をとったとしても、システムはユーザとは認識しない。

他者が最初にユーザとして特定された場合、または、発表者が交代する場合、ロックを解除する必要がある。ロック解除は以下の2つのどちらかで行う。

- ユーザが Kinect の視界から外れること。
- キーボードまたはマウスでシステム画面の「reset」のボタンを押すこと。

## 2.4 動作と音声による PC の操作

提案システムでは、ユーザの動作と音声によって PC を制御する。操作の多様性とユーザの負担軽減を考慮して、前者 (モーションコントロール) を、単純操作のみ対応する「姿勢モード」とメモ記録等を行える「マウスモード」で構成する。一方、後者 (音声コントロール) には前者の姿勢モードより機能を絞ってより単純な操作のみ対応させ、ユーザの発話の障害にならないようにする。なお、両者は同時に動作するが、後者は前者の姿勢モード内で起動・停止させる。図3にコントロールとモードの関係を示す。

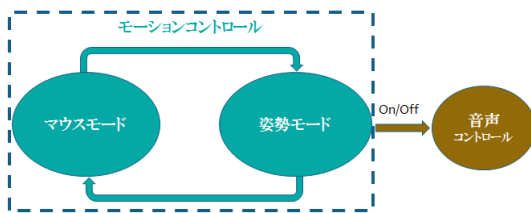


図 3: 操作モードの変更

## 2.5 モーションコントロール

- 姿勢モード (Single character mode)

姿勢モードは、対象アプリケーションの基本操作に対するモードである。ユーザの姿勢を認識して、対応付けられている PC のコマンドを実行する [3]。姿勢モードで対応するコマンドは、プレゼンテーション時にユーザが良く使うもので、マウスやキーボード等を使わないものとする。

ユーザをロックした直後に、ユーザの左手、右手、左肩、右肩と頭のスケルトントラッキングを取得する。そして、取ったスケルトントラッキングの中に、2つ以上のスケルトントラッキングの座標がある姿勢の条件を満たせば、コマンドを実行する。以下は姿勢モードの指令表である。

- マウスモード (Dragging mouse mode)

マウスモードでは、対象アプリケーション上で、マウス移動を伴う操作を実現する。ユーザの右手がマウスの動きをコントロールし、左手の上がりと下がりにはマウスの左クリックを表す。パワーポイントでメモ記録する時、右クリックは使わないため、提案システムでは、右クリックを表さない。アプリケーションがパワーポイントの場合、マウスモードはペン (描画) または消しゴム (消去) に相当する。プログラムの大まかな流れは、

表 1: 姿勢モードの指令表

PPT 操作	右手	左手
ページに進む	右に振る	-
前ページに戻す	-	左に振る
最初ページにリセット	-	上に上がる
他アプリケーションの切り換え	-	右肩をタッチする
マウスモード変更	上に上がる	-
音声認識起動	左肩をタッチする	-

1. マウスモード起動
2. 右手の位置の検索
3. 右手の位置の座標をコンピュータ内の座標に置き換える
4. 座標の位置にマウスの座標を移動
5. 姿勢モードに変更

ここで、姿勢モードに変更の指令を出るまで、2,3,4 を繰り返すような内容である。以下はマウスモードの指令表の指令表である。

表 2: マウスモードの指令表

PPT 操作	右手	左手
ペンを移動する	マウスを模擬し移動する	手を下がる
メモする/消す	マウスを模擬し移動する	頭の高さに上がる
ペンから消しゴムに変更する	-	頭より高い所に上がる
姿勢モードに変更する	-	左に振る

## 2.6 音声モード (Voice control mode)

音声モードでは、Windows システムで実行している対象アプリケーションのコマンドに単語を対応させる。音声モード起動中には、ユーザの話した単語を認識し、認識した単語と事前に保存した単語リストと比べ、一致した単語に対するコマンドを実行する。単語の認識については、Microsoft Speech SDK を利用する。この SDK を用いて、単語リストにある各単語に対する発せられた単語の近似率を求める。この近似率が設定した閾値と比べて、大きければ、対応するコマンドを実行する。以下は音声モードの指令表の指令表である。ここで、モーションコントロールのメモ記録以外の操作全部対応することになる。

## 3 システムの実装

### 3.1 実行環境

前章で述べた提案システムを実装する。Kinect SDK の性質上、提案システムの対応 OS は Windows 7 に限

表 3: 音声モードの指令表

PPT 操作	ページの進む	前ページの戻す	最初ページにリセット
音声	[Forward]	[Back]	[Reset]
PPT 操作	他アプリケーションに切り換え	メモ記録に変更	ペンと消しゴムの切り換え
音声	[Tab]	[Pen]	[Change]

定される。プログラミング言語にはC#を用いる。その他の実装環境を表4に示す。提案システムはWindows 7以上の一般的な性能であるPC、またはノートパソコンで実行可能である。

表 4: 実装環境

Operating System	Microsoft Windows 7 Professional SP1
CPU	Intel (R) Xeon (R) CPU X5450 @3.00 GHz 2.99GHz
メモリ	8.00GB
開発環境	Microsoft Visual Studio 2010 Professional
プログラミング言語	C#
Input device	Kinect for Windows[4]
Device driver	Kinect for Windows SDK v1.7

### 3.2 ユーザと各デバイスの配置

提案システムは、Kinect による直感的な操作を実現するため、ユーザの骨格情報を Kinect が認識できる場所に配置する必要がある。Kinect はユーザを邪魔しないように、部屋前方の高さ 1 m に設置し、プロジェクタと反対側で設置する。プロジェクタを見ながら操作を行うために、ユーザは Kinect がある部屋の前方を向く。後方の聴衆の方を向いて話をしたい場合には動作の左右は逆になる。

### 3.3 実行結果

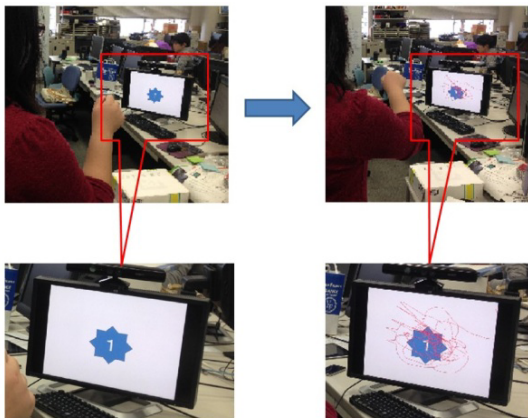


図 4: 実行結果

姿勢モードで、ユーザがPPTの最初ページに reset の動作とフローティングウィンドウが図 (5,6,7,8) で表す。

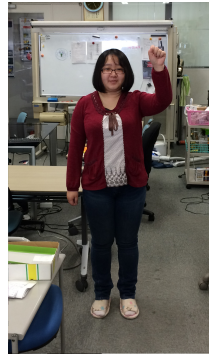


図 5: 最初ページに reset の姿勢

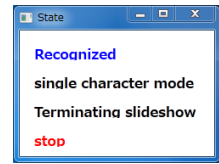


図 6: 最初ページに reset のフローティングウィンドウ

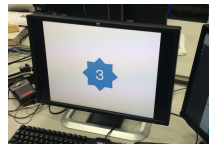


図 7: 最初ページに reset 前

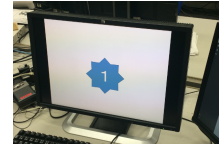


図 8: 最初ページに reset 後

## 4 評価と考察

### 4.1 評価

本システムの評価を行うために、提案システムに関する実験とアンケートを行った。実験は中央大学後楽園キャンパス 3 号館の 3501 室で実際に使用しているゼミ室で行う (図 9)。その後、被験者を対象として、システムの操作性や有用性に関するアンケートを行う。

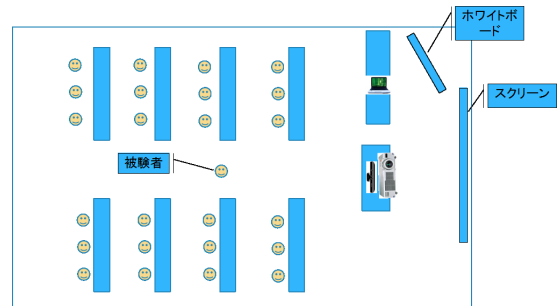


図 9: ゼミ室の図面

実験とアンケートは、日本国籍 13 人、中国国籍 9 人で合計 22 人を対象に実施した。この中で Kinect の経験者 6 人、経験無しの方は 16 人である。合計 30 問で、「ユーザ認識」「モーションコントロール」「音声コントロール」と「提案システムの有用性」から、経験有無と国籍より、ジェスチャ・音声による操作について評価を考察する。

## 4.2 考査

### 1. ユーザ認識

ロック前にロック指示以外の動作が反応しないことと一人だけロックすることが表現出来たが、ロック指示を認識することについて、ロック指示を認識できないことが多い。そして、経験者は未経験者より自分の動作を調整することが上手なので、ユーザ認識の成功率が高いである。また、低評価の理由欄で、経験者1名から「Kinectとの距離を調整しないと認識できない」と未経験者2名から「立つより座るのが認識やすい」との記述があることから、Kinectの設置場所、特に高さの調整が必要と考えられる。

### 2. モーションコントロールの姿勢モード

「ページに進む」、「前ページに戻る」、「他のアプリケーションに切り換え」と「最初ページにリセット」については経験者と未経験者に関わらず、高評価を得られた。この4つの操作は簡単な姿勢で経験に関係なくうまくコントロールできると考えられる。しかし、姿勢モードから円滑にマウスモードに変更することができないため、マウスモード変更の動作を変更しないとはいけない。また、「音声認識 On」と「音声認識 Off」の成功率は被験者より違い、音声の On と Off のアルゴリズムの改善しなければいけないと考えられる。

### 3. モーションコントロールのマウスモード

「動作モードに変更」、「ペンから消しゴムに変更」と「消しゴムからペンに変更」については、経験者と未経験者に関わらず、高評価を得られた。しかし、未経験者の評価の分散が大きくなって、未経験者は10分間で提案システムのモーションコントロールを慣れる人が多いと考えられる。

### 4. 音声コントロール

音声コントロールは英語で操作するため、母語が英語ではない人に対して使い辛くなった。また、各質問の分散について、被験者の英語会話のレベルより、ある単語に対する発音の正解率が違う。特に日本人の学生では、「Forward」「Reset」と「Tab」を正しい発音できない人もいる。

### 5. 提案システムの有用性

ここでは、各設問の相関を求め、実現した機能よりシステム有用性について、設問の組み合わせ考察することである。計算した結果から、「アプリケーションの切り換え」と「メモ記録」の有用性についての設問と相関のアンケート設問の相関係数が0.7以上である、各機能に高い評価を与えた被験者は該当機能がプレゼンテーションで役に立つと思うことと考えられる。つまり、該当機能は実際にプレゼンテーションする時、役に立つと考えられる。そして、「音声コントロール」の有用性についての設問と正解率が高い項目の相関係数が0.6以上であり、成功率が高ければ、プレゼンテーションする時役に立てると思うことと考えられる。つまり、発表者の母語を対応しなければいけない。特に日本語である。

## 5 結論

### 5.1 研究の成果

本研究では、ハンズフリーの状態ですべての動作・音声よりPCのアプリケーションを操作し、教員の授業

支援システムを提案した。提案システムでは、Kinectを用いることで機材等の持ち物が無い。ユーザは何も持たずに、教室のどこでもPC端末のアプリケーションをコントロールすることができる。

アンケートによる評価の結果、実際の実験を行った経験、他の教育ツールと比べ、提案システムの有用だと思えるか、提案システムのモーションコントロールの操作が便利ですか、提案システムの音声コントロールの操作が便利ですか、という項目において、70%以上の被験者から高い評価を得られた。このことから、提案システムはプレゼンテーション、ゼミなど行っている人に有用なシステムだと考えられ、本研究の目的達成に適したシステムであるといえる。

## 6 今後の課題

今後の課題としては以下が挙げられ、これらを解決することで目標とするプレゼンテーション用PC操作支援システムのより高度な実現が達成できると考えられる。

#### ● 音声認識の言語対応

本研究の評価で実施したアンケート評価の結果より、音声認識に対して、英語の発音についての要求がかなり厳しいである。これから、他国の言語も追加する。今は日本語と中国語対応を考えている。

#### ● Kinect 初心者でも容易に操作できるの実現

本研究の評価で実施したアンケート評価の結果より、Kinect 初心者は最初からうまくコントロールできないため、姿勢は認識できず、長い時間で疲労感も感じやすいである。解決策としては、小さな動きでもユーザのジェスチャーを認識できるようにする、または姿勢を改善する。

#### ● 新たな Kinect 2.0 の対応

今年には Kinect 2.0 for Windows が販売予定である。現在のシステムは Kinect 2.0 for Windows に対応機能、または新たな Kinect には何か新しい技術を持つかを確認し、それに現在のシステムが有利な機能を検討する。

## 謝辞

本研究を通じ、懇切丁寧な御指導、御鞭撻、及び多くの御支援を賜りました、中央大学理工学部情報工学科 牧野光則教授に深く感謝致します。

## 参考文献

- [1] 紅林秀治, 小林健太, 高山大輝, “KINECT センサーを用いた簡易動作分析システムの開発”, 日本産業技術教育学会誌 55(3), 213-220, 2013
- [2] 高橋弘毅, 難波道弘, 宿院頼, “タブレット端末を用いた教育実践報告 (新時代の学習評価/一般)”, 日本教育工学会研究報告集 13(4), 47-52, 2013
- [3] 知っている便利な PowerPoint のショートカットキー  
<http://kedu.kenet.ed.jp/hp/tips/tips026.pdf>  
(最終アクセス 2014 年 2 月 14 日)
- [4] Microsoft, Kinect,  
<http://www.microsoft.com/en-us/kinectforwindows/>  
(最終アクセス 2014 年 2 月 14 日)