

プレゼンテーションにおけるジェスチャーによるスライド操作機構の試作 On a Gesture Recognition Mechanism for Realizing a Slide-Show Presentation

山田 裕之[†] 丹羽 佑輔[†] 白松 俊^{††} 大園 忠親^{††} 新谷 虎松^{††}

Hiroyuki Yamada, Yusuke Niwa, Shun Shiramatsu, Tadachika Ozono, Toramatsu Shintani

1 はじめに

プレゼンテーションのスライド操作方法としては、PCの前に立ち、そのPCに付随するマウスやキーボードで操作することが一般的である。さらに、近年ではプレゼンテーション用マウスを使用したスライドの遠隔操作や、タッチパネルを使用して自らの指を入力とした操作といった、多種多様な入力デバイスが利用されている。

プレゼンテーション時において、汎用的に使用されると考えられるスライド操作を3つに分けると、ページ操作、ポインタによる注目点の指示、実際のデモやデモ映像の再生が挙げられる。しかし、既存のデバイス単体の入力を用いてすべて行おうとすると、プレゼンテーションから一旦離れなければならないか、必要な入力回数が増えてしまうといった問題がある。そこで、ジェスチャー入力に注目した。ジェスチャー入力は、発表者は何もデバイスをもつ必要がないため、既存の入力デバイスと比較して直感的に操作できると考える。ところが、ジェスチャー入力ですべてを実装してしまうと、発表者は、操作で使用する多くのジェスチャーを覚えなければならない。よって、ジェスチャー入力ですべての操作を置き換えるのではなく、使い慣れた既存の入力デバイスの利点を生かしつつ、ジェスチャー入力を組み合わせることで問題の解決を図った。センサーデバイスにはMicrosoft社から発売されたKinectを使用した。Kinectは、映像だけでなく距離も測ることができる。そのため、PCから離れた場所からの操作は勿論のこと、暗い場所でのプレゼンテーションにおいてもKinectの認識範囲内(1.2m~3.5mの位置[1])であれば、どこからでもスライドを操作することが可能である。

本稿では、プレゼン用マウスとジェスチャー入力を組み合わせたスライド操作支援機構と、Kinectの距離センサーを用いてタッチ入力を可能にさせたスクリーンもしくはディスプレイ(以下、表示装置とする)とジェスチャー入力を組み合わせたスライド操作支援機構を提案する。

2 ジェスチャー入力によるスライド操作

2.1 スライド操作

今回想定している操作は、ページ操作、ポインタによる注目点の指示、実際のデモとデモ映像の再生の3つである。ページ操作は、ページ送り、ページ移動、ズームといった3つのアクションを、また、ポインタによる注目点の指示では、ポインティング、ポインタアイコンの変更といった2つのアクションを、実際のデモとデモ映像の再生では、アプリケーションの切り替え、音量調節といった2つのアクションでそれぞれ実現する。

2.2 ジェスチャー入力による操作指示方法

Kinectによるジェスチャー入力とプレゼンテーション用マウスもしくはタッチ入力を可能にさせた表示装置を組み合わせた操作指示方法を提案する。

操作指示の流れとしては、図1のようなGUIを表示装置の隅に設置する。GUI上の操作の種類(以下、モードとする)ごとに割り当てたボタンを、発表者はマウスおよびタッチパネル化した表示装置を使って選択する。ジェスチャー入力は、選

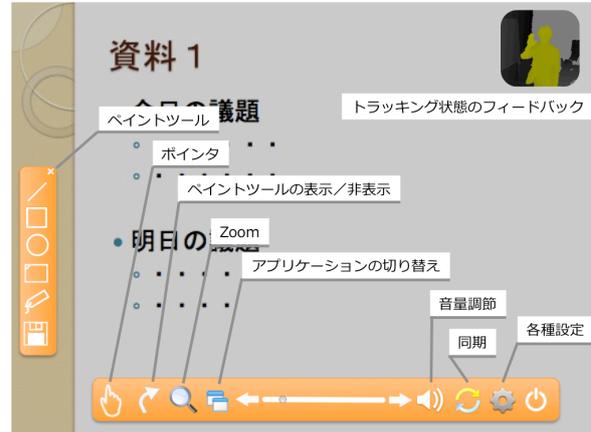


図1: タッチ入力を使用した場合のGUIの例

択したモードに沿ったものを行う。GUI上にはフィードバックとして、発表者のトラッキング状態とジェスチャー入力が成立したかどうかの情報を知らせる。

プレゼンテーション用マウスとジェスチャー入力を組み合わせる場合においては、プレゼンテーション用マウスの利点をそのまま活用する。プレゼンテーション用マウスの利点として、ページ操作に長けている点と精確にポインティングできる点が挙げられる。よって、ページ操作やポインティング等、1,2回程度の入力、ポインティング操作で行えるものはマウスに担当させる。ジェスチャー入力は、マウスによる操作を起点とした次なる処理の操作とした。例えば、スライドの拡大/縮小の操作を行う時、マウスのポインティングで拡大/縮小を行う中心点を決め、ジェスチャー入力によって拡大/縮小幅を決める。

タッチ入力を可能にさせた表示装置とジェスチャー入力を組み合わせた場合においては、タッチ入力でもモード(ポインティングモード、ズームモード等)を選択する。選択後の操作は、ジェスチャー入力で行う。操作の基準点を決めるようなポインティングをタッチで行おうとすると、発表時にスライド資料が腕で隠れてしまうため、ポインティングはジェスチャー入力で行う。この組み合わせによるプレゼンテーションでは、何ひとつデバイスを持つ必要がない。

スライド資料にグラフィックを表示させたくない場合として、ジェスチャー入力だけで操作可能なUIも試作した。こちらは、スライド操作をページ操作、ポインティングに絞っている。

3 システム構成

図2に示すように、本システムは、ユーザ検出・トラッキング機構、および、イベント推論機構から構成される。ユーザ検出・トラッキング機構は、Kinectから得られた情報を解析し、イベント推論機構のワーキングメモリへ渡す機構である。イベント推論機構は、ワーキングメモリに格納された情報とルールベースから、適切なイベントを選択し、発行する機構である。

[†] 名古屋工業大学 工学部 情報工学科

^{††} 名古屋工業大学大学院 工学研究科 情報工学専攻

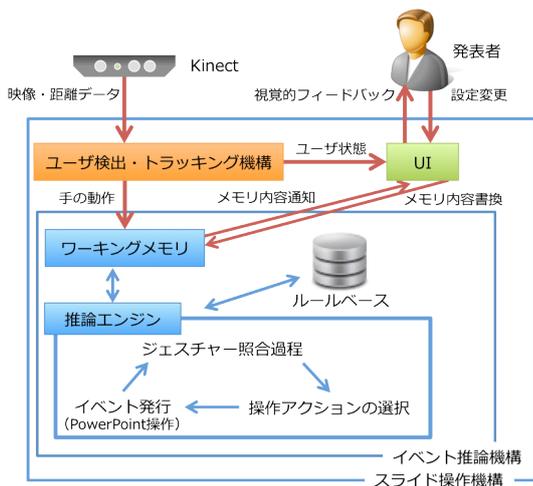


図 2: システム構成図

3.1 ユーザ検出・トラッキング機構

本機構では、Microsoft 社から提供されている KinectSDK を使用して、ユーザの検出と骨格のトラッキングを行っている。骨格トラッキングは、体の各部位ごとに分解された情報をもつ。トラッキング状態は、図 1 右上のように GUI 上に表示される。発表者はこの視覚的なフィードバックを通して、検出、トラッキング状態を確認できる。骨格トラッキング情報から、手の位置と距離の数値データを取得し、手の動作方向（左右上下、前後）を識別する。手の動作方向は、体から手がある一定の距離があいた時を起点とし、その点からどの方向に手が移動したかを識別する。GUI 上には、手の起点や識別された手の動作（左右上下、前後）といった一連の状態が表示される。手の動作方向および左右右手の情報は、イベント推論機構のワーキングメモリへ格納される。

3.2 イベント推論機構

イベント推論機構は、ルールベースシステムを用いて実装した。ワーキングメモリには、手の動作情報と現在のモード状態が格納されている。現在のモード状態とは、選択されたボタンが指すモードであり、手の動作が何の操作を目的としているのかを表している。モードが何も選択されていない場合に行われる手の動作は、すべてページ操作のための動作としている。ルールベースには、手の動作の組み合わせからジェスチャーを決定するルールが格納されている。ルールの条件部には、ジェスチャーを決定するためのいくつかの手の動作組み合わせが記述されている。実行部には、ワーキングメモリ内の情報の消去あるいは次の操作のトリガーとなる情報の追加、変更および、スライドを操作するためのイベントが記述されている。これらを推論エンジンを使用して、ワーキングメモリとルールベースとを照らし合わせる。もし、マッチしたルールがあれば、ルールの実行部に従ったイベントを発行する。発行されたイベントは、プレゼンテーションソフトに送られ、スライド操作が実行される。

3.3 表示装置のタッチパネル化

表示装置をタッチパネル化するにあたっての処理の流れを説明する。(図 3 参照) まず、Kinect から得られた映像データから、表示装置の頂点をユーザが選択する。選択した 4 点からなる四角形で形成される図形を本来の表示装置の 4 隅の座標と対応させる。画像認識に関連する機能のライブラリである OpenCV を用いて射影変換行列を求める。次に、その変換行列を使用して、ボタンオブジェクトの領域を、カメラから見た表示装置の座標系に再配置する。発表者の手が、変換されたボタン領域内にきた場合、同座標点における表示装置と手の深度情報を比較して、ボタンに接したかどうかを判断する。

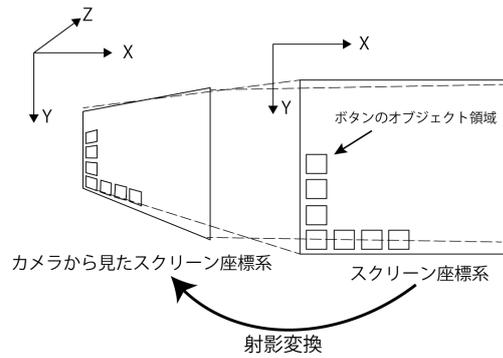


図 3: カメラから見た表示装置の座標系への変換

4 システムの評価と考察

本システムを使用したプレゼンテーションでは、本来では目的の処理までに多くの操作が必要であったものを、既存のデバイスによる数回程度の入力と簡単なジェスチャー入力だけで操作できるようになった。認識精度は、Kinect 正面に向けてのジェスチャー入力においては良好である。

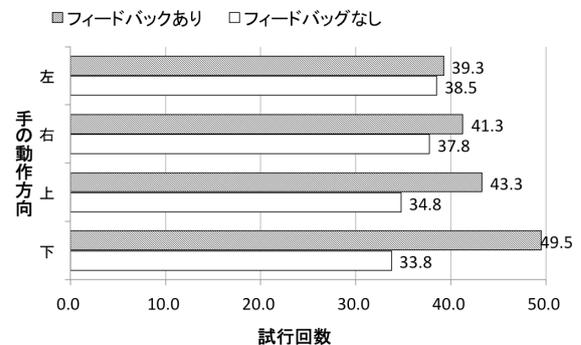


図 4: 評価グラフ

ここでは、ユーザ検出・トラッキング状態のフィードバックを GUI 上に表示する必要性を評価した。4 人の被験者に協力してもらい、フィードバックがある場合とない場合で簡単なスライド操作の認識回数を評価した。50 回ずつ順に左右上下（計 200 回）に腕を動かしてもらい、その操作の認識回数の平均を比較した。フィードバックには、起点が決められたという情報と、そこから腕が左右上下どちらに動いたかの情報を含めた。Kinect から被験者までの距離は 2m で固定した。結果を、図 4 に示す。フィードバックがある場合の方が認識率が高いことがわかる。

5 おわりに

プレゼンテーションにおける自然な入力として、指示棒を入力デバイスとしたものがある [2]。指示棒の先端を認識させることで、スクリーンへの入力を行っている。このシステムは、指示棒のみを入力としているため、多くの操作はできない。しかし、タッチ入力可能な表示装置およびジェスチャー入力を組み合わせることで、より複雑な操作を行うことが可能である。本稿では、ジェスチャー入力によるスライドの直感的な操作を実現するための、ジェスチャー入力と既存のデバイスを組み合わせる直感的な操作手法について述べた。使い慣れた既存のデバイスとジェスチャー入力を組み合わせることで、少ない入力でもより直感的に、スライドを操作することができるシステムを設計、構築することができた。

参考文献

- [1] Microsoft Research, Programming Guide: Getting Started with the Kinect for Windows SDK Beta, 2011.
- [2] 佐藤 周平, 柴山 悦哉, 高橋 伸: “指示棒の認識を用いたプレゼンテーションシステムの構築”, コンピュータソフトウェア 17(3), 269-272, 2000.