5ZC-1

スクリーン仮想タッチを実現した プレゼンテーション支援システムとそのセットアップの効率化

山田 裕之 † 白松 俊 †† 大囿 忠親 †† 新谷 虎松 ††
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 ††

1 はじめに

本研究では，プレゼンテーションにおけるスライドの操作をジェスチャー，もしくはスクリーンへのタッチで行えるインターフェイスを作成した．ここでは，センサーデバイスである Kinect 1 台だけを用いて，ジェスチャー認識とスクリーンへのタッチ入力を実現する．大型ディスプレイにおいても，タッチ入力を利用可能にするために，ジェスチャー認識で使用する Kinect をタッチの検出センサーとしても使用し，スクリーンを仮想的にタッチ検出可能にする．加えて，スクリーン仮想タッチを実現する環境を構築するためのセットアップの効率化を図る．本システムを用いたスライドの操作は，物理的なデバイスを持ったり，身につけたりする必要はない．ジェスチャー入力と他の入力手法を組み合わせた研究も行った [1].

本稿では，ジェスチャー入力とタッチ入力によるスライドの操作指示方法，スクリーン仮想タッチの実現方法，および本システムに対する考察を述べる．

2 スライドの操作指示方法

本システムでは，発表者が行えるいくつかの操作を表すものをモードと呼ぶ．例えば，ページ操作モードでは，スライドのページめくりが行える．拡大縮小モードでは，ページの拡大縮小が行える．ジェスチャーは，モードごとに割り当てられたジェスチャーと，モード中での操作用ジェスチャーがある．発表者は，モードをスライドの操作種に合わせて選択後，スライドの操作を決定する．発表者が覚えなければならないジェスチャーは，モード数と操作数に依存する．すなわち，より多くの操作を実現するほどユーザビリティが低下する．そこで，モード選択の補助的なツールとしてスクリーン仮想タッチによるモード選択機能を追加した．スクリーン仮想タッチとは，発表者がスクリーンに触れることができる場合，スクリーンをたタッチすることでモードを選択できる機能である．スクリーンへのタッチの検出は，各モードが割り当てられた領域（以下，モード領域と呼ぶ）に発表者の手があるかどうかで判定する．各モードがスクリーンのどこにモード領域を持つか事前に知っている必要がある．モード領域を視覚的に伝えるために，図 1 で示すような各モードを表すボタンアイコンを表示する機能も追加した．これらのボタンアイコンは，発表者の手がスクリーン領域内に存在する時に表示される．ボタンアイコンは表示／非表示の切り替えができる．

スクリーン上には，フィードバックを表示させることも可能である．フィードバックは，骨格の追跡状況，現在のモード状態，ジェスチャーが適正に処理されたかの通知を示す．フィードバックの表示の有無は切り替えることが可能である．　発表者のジェスチャーとしては，反映されるスライドの操作を全く連想できないジェスチャーや，聴講者に余計な



図 1: モード選択用ボタンアイコン

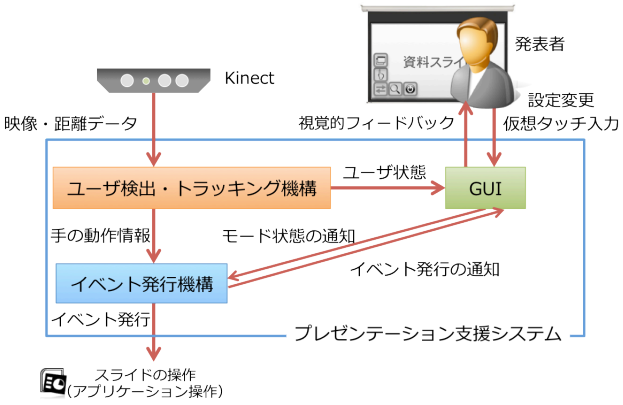


図 2: システム構成図

疑問を抱かせるジェスチャーは避けるべきである．本システムのジェスチャー入力では，スマートフォンやトラックパッドでよく利用されている操作を参考にジェスチャーの動作を決めた．例えば，ページ操作モードでは，指を一定方向へ滑らせるスワイプを参考にして，手を左右に振る動作で前ページへ戻る，次ページへ進む．拡大／縮小モードでは，2 本の指で操作対象をつまむように動かすピンチインアウトを参考にして，両手を前に突き出し，左手と右手間の距離を広めたり狭めたりすることで拡大，縮小率を変更する．

3 プレゼンテーション支援システムの構成

プレゼンテーション支援システムの実装方法について述べる．システム構成図を図 2 に示す．

ユーザ検出・トラッキング機構では，発表者の検出と骨格追跡に，手を振った方向の判別を行う．発表者の検出と骨格追跡に関しては，KinectSDK を利用する．この時得られる手の位置情報から算出した手の移動量を元に手の動作方向を判別する．イベント発行機構では，手の動作方向の判別結果からスライドを操作するためのイベントを発行する．

手の動作方向の判別方法について述べる．手の動作方向

Presentation Support System based on Virtual Touch Action and Streamlining System Setup

Hiroyuki YAMADA, Shun SHIRAMATSU, Tadachika OZONO, and Toramatsu SHINTANI

†Dept. of Computer Science, Nagoya Institute of Technology. ††Dept. of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology.

Copyright ©2012 Information Processing Society of Japan.
All Rights Reserved.

は，手の移動量が閾値を超えたか否かで判別する．Kinectとユーザ間の距離が大きいほど，手の移動量は小さくなる．したがって，距離を考慮に入れて閾値の補正を行い，判別精度の向上を図った．事前に，ユーザに一定の距離ごとに計測した各動作方向に対する手の移動量を計測し，Kinectから0.5mおきに各動作方向に対する手の移動量とあわせて記録した本データを，移動量基礎データと呼ぶ．本データより，距離を考慮にいれた閾値を決定する関数を定めた．

$$\theta = [\ x^n \quad x^{n-1} \quad \cdots \quad x \quad 1\ ]A^{-1}B$$

ここで，xはKinectからユーザまでの距離である．行列A,Bは学習データである．行列Aは手の移動量，行列Bは行列Aの各要素に対する計測時のKinectからユーザまでの距離である．A-1Bは，閾値θを決める関数の係数となる．

次に，スクリーン仮想タッチの実現方法について述べる．仮想タッチ入力の実現には，2つのフェーズ，キャリブレーションとタッチ入力判定がある．タッチ入力判定は，ユーザの手がある1つのボタンアイコンをタッチするような動作をしたかどうかで判定する．すなわち，ユーザの手の位置座標がモード領域内に存在するかしないかが重要である．しかし，ここでの手の位置座標は，Kinectがもつ座標系における位置である．加えて，Kinectから見たモード領域は，正方形でも長方形でもない四角形である．したがって，キャリブレーションを行い，Kinectから見たときのスクリーン上にある各モード領域の正確な位置座標を算出する必要がある．最初は，スクリーンの検出である．スクリーンの検出は，スクリーンに決まった色を表示し，そこからスクリーンの範囲を検出する．具体的に説明すると，Kinectからの映像よりHSVを用いて決まった色を抽出し，2値化する．2値化した画像からエッジ検出を行い，スクリーンの周りの辺を検出する．次に，検出したエッジからHough変換より直線を検出し，スクリーンの4隅の座標点を算出する．この4つの座標点に対応する，クライアント座標系におけるGUIの4隅の座標点から，OpenCVを利用し，射影変換行列を算出する．求めた射影変換行列を，GUI上のモード領域に適用することで，求めたいKinectから見たスクリーン上のモード領域を得られる．この得られた領域内にユーザの手の位置座標があるかどうか内外判定を行うことで，スクリーン仮想タッチが実現できる．この実現方法より，本システムを利用するにあたってのセットアップは，Kinect1台を適切な位置に設置するだけの効率的なセットアップとなる．

4 考察

スクリーン上にユーザへのフィードバックを通知する機能があるが，これらのフィードバックは2つの重要な役割をもつ．1つは，本システムが認識できない誤ったジェスチャーをユーザが行った場合，認識できなかった事実を伝えること．もう1つは，操作方法をユーザに学ばせることである．フィードバックがある場合，ユーザはジェスチャーが思い通りに認識されたかどうか確認できるため，表示していない時と比べて，誤って認識されるようなジェスチャーを繰り返し続けることを防げる．　本研究では，ジェスチャー入力とタッチ入力を組み合わせたスライドの操作を提案した．ジェスチャー入力とタッチ入力は，共に物理的なデバイスを持ったり身につけたりする必要がない．タッチ入力をモードの選択，ジェスチャー入力をスライドの操作と使い分けることでユーザが覚えなければならないジェスチャーの数を減らし，ユーザビリティを向上した．スライドの操作をジェスチャーだけで実現した場合と比べ，単純なジェスチャーを各モードで使い回すことでより多くの操作が可能となる．これらのことから，ジェスチャー入力とタッチ入力を組み合わせたスライドの操作は利便性が高いといえる．ユーザのスライド操作時の負担の削減方法としては，ユーザの骨格情報を利用するという方法も考えられる．KinectSDKは身体の20もの部位の位置情報を取得できる．これら骨格情報を用いて，プレゼンテーション中に発表者が資料の前に立っても，遮蔽されないように投影するオブジェクトを動的に配置する研究がある[2]．Kinectを用いたスライドの操作方法には多くの拡張性があり，身体全体を使ったスライドの操作といったものも考えられる．

一般的に，ジェスチャーを認識するための環境は，カメラを複数台用いて，念入りな調節の元で構築する．本システムでは，Kinect1台を適切な位置に設置するだけで実現できるため，セットアップに要する時間を削減できるメリットがある．しかし，Kinect 1台だけを用いた場合のデメリットも存在する．Kinectのユーザ検出可能な有効範囲は，Kinectからの距離が1.2m〜3.5mである．演壇や教壇で，本システムを利用する場合，発表者はKinectの認識可能範囲内に存在しなければならない．認識範囲の広さの面では,Kinect複数台を用いた方が有用であるといえる．また，ジェスチャーの認識精度向上のために，ジェスチャー認識はSVMを用いたパターン認識手法など，多様な手法が提案されている．これら既存手法と本研究で使用している手法を比較し，さらなる精度向上を行うことが可能である．

スクリーン仮想タッチに関しては，発表者はスクリーンの方向を強く意識して操作する必要がある．しかしながら,Kinectを用いた場合ではスクリーンに必ずしもタッチする必要はない．空中に領域を設定し，その領域に発表者の手が入ったとき，タッチとみなす手法も考えられる．

聴講者にモード切り替えをしようとする様子が明示的に伝わってしまうといったプレゼンテーションの流れを阻害する問題を内包しており，今後の研究で考慮する必要がある．解決策として，タッチ入力だけでなく他の入力手法とも組み合わせ，それらを比較する必要があると考える．

視覚的フィードバックに関しては，発表者のジェスチャー入力の精度を高めるものであるが，プレゼンテーションの質を落としてしまう可能性もある．そこで，発表者のジェスチャー入力の精度向上のために，発表中に支援するのではなく，事前に学習してもらう方法も考えられる．例えば，ジェスチャー入力によるスライドの操作手法に対してのチュートリアル機能を，本システムに備えるなどがある．

5 おわりに

本研究では，ジェスチャー入力とタッチ入力を組み合わせたインターフェイスを作成した．発表者がスライドを操作するために覚えなければならないジェスチャーの数を減らすことで，ユーザビリティを向上した．単純なジェスチャーだけでより多くの操作を実現できる．また，本システムを誰もが手軽に利用できるように工夫を施した．本システムでは，通常のスクリーンをタッチ検出可能な表示デバイスにする．よって，ユーザがタッチ検出可能な表示デバイスを用意する必要はない．ジェスチャー認識とスクリーン仮想タッチをKinect1台だけで実現した．ユーザは，Kinectを適切な位置に設置するだけで，本システムを利用することが可能である．

参考文献

[1] 山田裕之，丹羽祐輔，白松俊，大囿忠親，新谷虎松：“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”，第10科学技術フォーラム論文集，pp. 529–530，Sep，2011.

[2] 前田晴己，黒澤祐也，栗原一貴，宮下芳明: “MAEDE:スクリーン前でのプレゼンテーションスタイル”，WISS2011，第19回インタラクティブシステムとソフトウェアに関するワークショップ論文集，pp.164-166，2011.

Copyright ©2012 Information Processing Society of Japan.
All Rights Reserved.