

# 頭の傾きと音声を用いたグラフィカルユーザインタフェースの試作

日大生産工 (学部)    ○中山 智晶    日大生産工 岡 哲資

## 1 まえがき

グラフィカルユーザインタフェース(GUI)は、パソコン、タブレット、スマートフォンなどのさまざまなコンピュータで幅広く利用されている。GUIの入力は、通常は、マウスなどの間接ポインティングデバイスやタッチスクリーンのタッチ・ジェスチャ入力など手を用いて行う。一方、手を使わないコンピュータの操作方法には、視線<sup>1) 2)</sup>、音声<sup>3)</sup>、脳波など、様々なものがある。これらは、障害者向けのユーザインタフェース、手が使えない状況、手が使いづらい状況での健常者向けユーザインタフェースへの応用が期待される。しかし、GUIへの応用を考えると、手を用いない入力方法は、手を用いる入力方法と比べると、入力に要する時間、ユーザの疲労、操作方法の難しさなどの問題が多い。

手を使わずに、視線、音声とウィンクを組み合わせるユーザインタフェースの研究も行われている<sup>4)</sup>。しかし、このインタフェースは、ウィンクが苦手な人には適さない。そこで、我々は、音声と頭の傾きを組み合わせたインタフェースを試作した。本稿では、試作したインタフェースについて報告する。

## 2 インタフェース操作の概要

試作したインタフェースでは、頭の傾きと音声を組み合わせてマウスカーソルの操作を行う。本インタフェースを用いると、手を使わずにマウスカーソルの移動、クリック、右クリック、ドラッグ、スクロールの各操作が行える。頭の傾きによって、マウスカーソルの移動方向、ドラッグ、スクロールの方向を指定する。音声によって、指定した場所へのマウスカーソルの移動、マウスカーソルの移動速度の変更、クリックなどを行う。

## 3 マウスカーソルの移動

### 3.1 頭の傾きによるカーソル移動

頭を右または左に 15 度以上回転させると、

マウスカーソルは、右または左に移動し始める。頭を元に戻すと、カーソルは停止する。また、頭を前または後に 15 度以上傾けると、カーソルは上または下に移動し始め、戻すと停止する。15 度以上の左右の回転（ヨー）と前後の傾き（ピッチ）が組み合わせると、カーソルは斜めに移動する。

### 3.2 カーソルの移動速度の変更

「スピード1」、「スピード2」、「スピード3」、「スピード4」、「スピード5」のいずれかの音声入力を行うと、カーソルの移動速度が変わる。また、「アップ」と入力すると、移動速度が1段階上がり、「ダウン」と入力すると、移動速度が1段階下がる。

### 3.3 音声によるカーソル移動

「左上」、「上」、「右上」、「左」、「右」、「左下」、「下」、「右下」、「ホーム」のいずれかの音声入力を行うと、カーソルが瞬時に画面内の特定の位置に移動する。図1に、ウィンドウサイズが 1280×960[pixel]の場合の移動先を示す。「ホーム」と入力すると、カーソルは画面の中心に移動する。

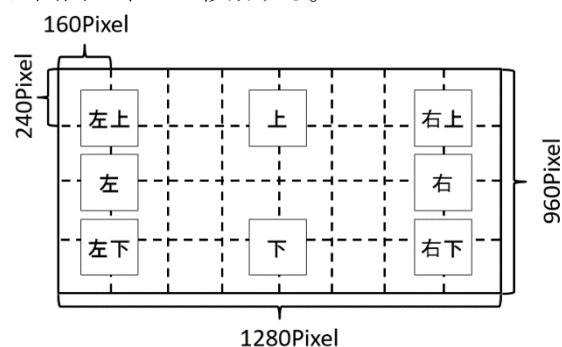


図1 音声によるカーソル移動の移動先

## 4 クリックなどの操作

音声で「クリック」と入力すると、左クリックを1回行う。「右クリック」と入力すると、

右クリックを1回行う。「ダブルクリック」と入力すると、左クリックを2回行う。

## 5 ドラッグとスクロール

音声で「ドラッグ」と入力し、カーソルの移動を行うと、ドラッグ操作が出来る。ドラッグアンドドロップを行うときは、まず、マウスカーソルを移動し、ドラッグしたいオブジェクトにマウスカーソルを合わせてから、音声で「ドラッグ」と入力する。次に、ドロップしたい場所へマウスカーソルを移動する。最後に、音声で「離す」と入力する。画面のスクロールも先述のドラッグと同様に、音声で「スクロール」と入力し、カーソルの移動を行う。音声で「離す」と入力すると、スクロールが終了する。ドラッグ中に音声で「スクロール」と入力しても無視される。同様に、スクロール中に音声で「ドラッグ」を入力しても、ドラッグ操作はできない。

## 6 使用機材とソフトウェア

本インタフェースは、RGB-D カメラとマイクとして Kinect v2 を使用した。また、マウスカーソルの操作を行う為のパソコン、頭の角度を取得するために Kinect for Windows SDK も使用した。図2のように Kinect を設置することでインタフェースを使用する。使用するパソコンの仕様を表1に示す。音声認識の為に、Microsoft Speech Platform SDK を使用した。

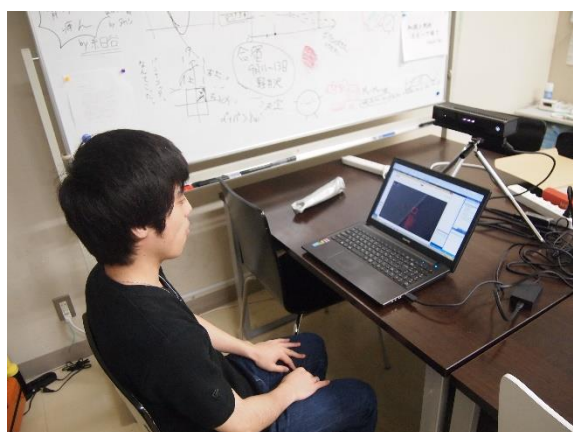


図2 試作したインタフェース

表1 パソコンの仕様

スペック	
オペレーティングシステム	Windows8.1 64 ビット
プロセッサ	Intel(R)Core i7-4710MQCPU@250MHz
実装 RAM	16.0MHz
グラフィックチップ	NVIDIA GeForce CTX850M

## 7 イベント処理

Kinect for Windows SDK によって算出された頭の回転角度 Yaw[degree]、Pitch[degree] から、式(1)、式(2)を用いて、マウスカーソルの座標である整数型変数 pX、pY に代入することで、マウスカーソルの移動を行う。この時、整数型変数 Speed には、移動速度に応じて 1, 2, 3, 5, 10 の値が代入される。

$$pX = pX + \frac{Pitch}{Speed} \quad (1)$$

$$pY = pY + \frac{Yaw}{Speed} \quad (2)$$

同様に、Microsoft Speech Platform SDK によって音声の入力内容を処理し、操作を行う。

## 参考文献

- 1) 岡本龍太郎、古本淳也、松原俊一、Martin J.Durst : 視線情報を用いた Web ブラウジング機能の提案と実装、情報処理学会第 77 回全国大会講演論文、No1(2015)、pp. 217-219
- 2) 真鍋宏幸 : ヘッドホン型視線入力インタフェース、精密工学会誌(2016)、82.1: p.7-10.
- 3) 世古拓海、河崎卓也、田村哲嗣、速水悟:実環境におけるマルチモーダル音声インターフェースの適用 (パターン認識・メディア理解)、電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, 113(493), p. 185-190.
- 4) 赤澤 恒 : 視線・ウィンクと音声を用いたユーザインタフェース、平成 28 年度日本大学生産工学部数理情報工学科卒業論文(2017)