

## 距離画像センサと音声認識を用いたヒューマンインタフェースに関する研究

Research of Human Interface using Image Distance Sensor and Speech Recognition Technology

武田 実十  
Minoru Takeda

新藤 義昭†  
Yoshiaki Shindo

### 1. はじめに

PC を操作するためのキーボードとマウスは、使用するソフトウェアによっては複雑な操作を要求される時がある。また、キーボードとマウスを置くための場所と環境が必要である。健常者には使いやすいインタフェースであるが、高齢者や障害者の多くはソフトウェアごとに異なる操作を覚えるのに多くの時間を必要としたり、そもそもマウスやキーボードを操作できない場合もある。この問題を改善するため、キーボードとマウスの代わりに、距離画像センサと音声認識を用いた直感的なヒューマンインタフェース方式を模索する。

### 2. 研究の目的

本研究の目的は、パーソナルコンピュータと操作者の間のヒューマンインタフェース方式として、距離画像センサと音声認識技術を用いた新たな対話型インタフェース方式を実現するための基礎研究を行うことである。研究の中核要素は、人間の手を距離画像センサで探知して、指操作をコマンド認識する技術を開発することである。指操作だけでは安定したヒューマンインタフェースが期待できないので、音声認識技術を併用する方針である。キーボードやマウスをうまく操作できない高齢者や障害者がコンピュータを操作する一方式としての可能性を探ることが目標である。

### 3. 研究の内容

操作方式の基礎研究の土台として、ペイントソフト（絵描きソフト）を題材として実験する。本ソフトウェアを FingerPaint と名づける。以下に、FingerPaint の設計仕様を示す。

- ① 作図開始指令（図形の種類）を音声で行う。
- ② 操作に対する大きさや変形等の操作パラメータを指の動きで指示する。
- ③ 指画像認識には、あらかじめ作成し登録しておいた照合用データベースを用いる。

例として、楕円を描く操作を以下に示す。

- (1) 操作者はウインドウ上に円を描くため、音声で「円を描いて」と発声する。これが作図開始指令コマンドとなる。
- (2) FingerPaint は、音声合成で、「どこに？」と操作者に尋ねる。
- (3) 操作者は指操作で位置を指定する。FingerPaint は、指操作をリアルタイムで認識して既定の大きさの円を描くが、指を動かすと同期して円の位置もリアルタイムに変わる。
- (4) その位置でよい場合は、操作者は、「オーケー」と発声する。
- (5) 次に、FingerPaint は、音声合成で、「どのくらいの大きさで？」と操作者に尋ねる。
- (6) 操作者は、指操作で大きさを指示する。
  - ◎ 指を画面に近づけると円は大きくなる。
  - ◎ 指を画面から遠ざけると円は小さくなる。
  - ◎ 指を上を動かすと円は垂直方向に伸びる。
  - ◎ 指を下を動かすと円は垂直方向に縮む。
  - ◎ 指を右を動かすと円は横方向に伸びる。
  - ◎ 指を左を動かすと円は横方向に縮む。

この操作により、円はリアルタイムに楕円として自由な大きさになる。

- (7) その位置でよい場合は、操作者は、「オーケー」と発声する。
- (8) 次に、FingerPaint は、音声合成で「角度はどうする？」と操作者に尋ねる。
- (9) 操作者は、指操作で角度を指示する。
  - ◎ 指を右に回すと図形も右回転する。
  - ◎ 指を左に回すと図形も左回転する。

この操作もリアルタイムで図形が回転する。

- (10) その角度でよい場合は、操作者は、「オーケー」と発声する。

†日本工業大学大学院 工学研究科 情報工学専攻  
Graduate School of Computer and Information Major, Nippon  
Institute of Technology

#### 4. システム構成

FingerPaint のシステム構成図を図 1 に示す。

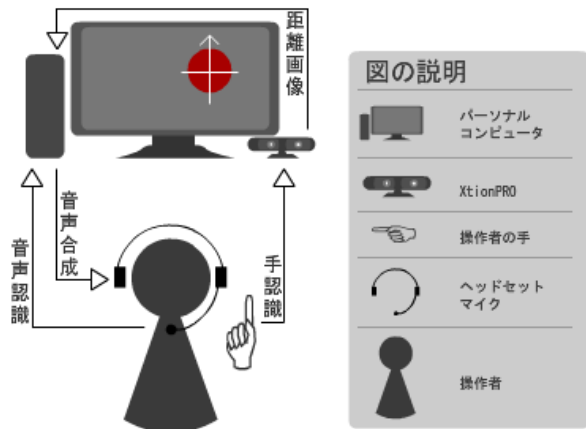


図 1 FingerPaint のシステム構成図

FingerPaint は、最初に操作者のヘッドセットマイクから入力された音声の認識を行い操作命令を開始する。次に XtionPRO を用いて操作者の手の距離画像を分析して、PC のディスプレイで手の動きに応じた図形を描画する。

#### 5. 距離画像センサ XtionPRO

XtionPRO は、ASUS が PC で使用することを前提に開発した距離画像センサカメラデバイスである。XtionPRO は、距離深度センサーと専用ソフトウェアを動作させるプロセッサを内蔵した 2 眼カメラであり撮影したものの動きをリアルタイムで認識することができる。本研究では、この XtionPRO から検出した操作者の手の動きを分析して、数値パラメータ（位置、大きさ、角度等）を入力する。

#### 6. 音声認識 Microsoft Speech API

本研究では、マイクロソフト社が音声認識及び音声合成エンジン用の標準インタフェースとして規格化している Microsoft Speech API と、日本語の認識用に SpeechSDK51LangPack を使用する。

#### 7. FingerPaint の動作画面

本研究で評価用に開発したペイントソフトウェアである FingerPaint の動作画面を図 2 に示す。操作者は、指操作と音声で画面上に色々な図形（線、円、多角形、特殊図形等）を組み合わせて表示させ、お絵かきをするシステムである。

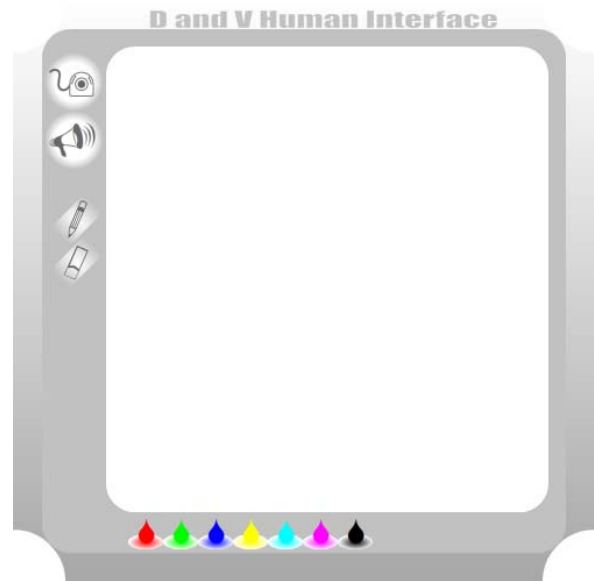


図 2 FingerPaint の動作画面

#### 8. 評価の方法

PC の操作にマウスやキーボードを用いるのが困難な障害者と高齢者を対象とした FingerPaint の操作実験を計画している。実験に際しては、あらかじめ用意した操作パターンで練習し、その後実際に使用してもらい、操作性や誤認識の評価を行う予定である。最後にアンケートを実施し有効性の評価を行う計画である。

#### 9. 参考文献

- [1] Introducing OpenNI  
<http://www.openni.org/>
- [2] モーションキャプチャデバイス XtionPRO  
[http://www.unitycorp.co.jp/asus/option/xtion\\_pro/](http://www.unitycorp.co.jp/asus/option/xtion_pro/)
- [3] 中島 淑貴: “新たなユニバーサルデザイン・入力インターフェースとして”, 第 12 回 ISID 学生論文 IT“夢”大賞
- [4] 龍淵信, 佐藤大輔: “カーナビ操作において入力デバイスの違いがもたらす負荷の比較検証”, ケータイカーナビの利用性と人間工学シンポジウム, 2002
- [5] 平野 緑, 大久保義美, 阪 美里, “パソコン苦手意識の世代差”, 日本教育心理学会総会発表論文集 (44), 424, 2002-08-09
- [6] 山口 亨, 櫻井 翔一郎, 佐藤英理, “指差し動作による人間-ロボット間のナチュラルインターフェース”, 22nd Fuzzy System Symposium Sapporo, Sept. 6-8, 2006