

音声認識とジェスチャ機能を用いた料理者支援 Cooking Assistance by Speech Recognition and Gesture Function

杉本 拓也†
Takuya SUGIMOTO

伊與田 光宏†
Mitsuhiro IYODA

1. はじめに

近年、大学生の一人暮らしなどで初めて料理をするユーザーが増えている。また、初めて作る料理のレシピをスマートフォンやPCなどを見ながら料理することが一般的なものとなっている。しかし、料理中は食材を扱うため、手が汚れてしまっていたり、濡れていることが考えられる。したがって、手が汚れてしまっていたり、濡れている状態でスマートフォンやPCを扱うと、スマートフォンのタッチパネルが傷ついたり、マウスやキーボードの故障が考えられる。そこで、PCをマウスやキーボードを使わずに操作することで料理のレシピを閲覧できるシステムがあれば便利であると考えた。

2. 概要

本研究では、Microsoft社から発売されている体感型インタフェース、Kinectを用いる。

Kinectを用いる利点は、PC操作を行う際にKinectは動きの認識に映像を用いるため、マウス等のデバイスを持つ必要がないこと、マイクを備えているため音声認識機能を使用できることが考えられる。したがって、ユーザーがPCから離れていたたり包丁などの調理器具を持っていることが予想される、料理中でもPCの操作を行うことができる。

3. ジェスチャ機能

ジェスチャ機能とは、手や腕などの身体の動きや体勢を読みとって何らかの動作を行う機能のことをいう。ジェスチャ入力の特長は、自然に操作を行うことができる可能性を秘めているという点である。本研究では、以下の操作をPCに入力する際にジェスチャ機能を用いる。

- 右手の動きに合わせてマウスカーソルを移動する機能
- 右手を2秒間停止させた状態にすることでマウスの左クリックを行う機能
- 左手を左に振るとマウスのダブルクリックを行う機能
- 左手を右に振るとマウスの右クリックを行う機能

4. 音声認識機能

音声認識とは、ユーザーが発する言葉を認識することによって文字情報として入力したりする機能を指す。しかし、認識する単語の数を増やし過ぎると、誤検出の可能性が増すため注意が必要である。本研究では、以下の操作をPCに入力する際に音声認識機能を用いる。

- データベースに登録されている料理メニューの入力を行う機能

- Webブラウザやスタートメニューを起動する機能
- 「次のページ」、「前のページ」、「閉じる」の入力をPCに行う機能

5. 事前検証

Kinectを用いたPC操作の利便性を検証する為に事前検証を行った。アンケート結果を表1として以下に示す。アンケートは5段階評価で表してもらい、1が「とても悪い」、5が「とても良い」とした。また、自由記述欄に意見を頂いた。

表1. アンケート結果

質問No.	質問内容	評価
1	ジェスチャは使いやすかったか	3.1
2	音声認識は使いやすかったか	4.2
3	誤認識はなかったか	2.3
4	素早く操作を入力できたか	2.7
5	操作性は良かったか	3.2

表1をみると、誤認識についての質問の評価が悪いことがわかる。これは、日常的にジェスチャと音声認識機能を使っていないユーザーの多くが、誤認識が多いと感じたという評価の1や2といった評価値をつけていた。一方で、ジェスチャと音声認識機能を使っているユーザーは、あまり誤認識があるとは感じなかったという評価の4や5の評価値をつけていたことがアンケート用紙から見てとれた。したがって、操作に慣れが必要になってしまったことがわかった。

6. システム

6.1 システム提案

事前検証を踏まえて、Kinectを用いてPCを操作するに当たって以下のような機能を付加することが望ましいと考えた。

- ジェスチャ入力によるマウスカーソルの移動、各種クリック操作を行う機能
- 音声入力によるWEBブラウザなどの呼び出しや料理名を入力する機能

ジェスチャ機能に関しては、前述したように手に調理器具を持っていても問題なくPCに操作を入力できるものとした。また、事前検証の結果から入力するジェスチャは学習機能を持たせることで誤認識の緩和につながると考える。音声認識機能に関しては、登録する単語をデータベース化し、自由度を持たせることでユーザビリティの向上を期待できるとともに登録単語の増加から生じる誤検出の緩和につながると考える。

6.2 システム構成

KinectとPCの接続構成は以下に示す図1のようになる。構成としては、Kinectによって得られた情報をPCに伝達し、伝達された情報が料理

† 千葉工業大学

Chiba Institute of Technology

メニューであった場合、料理データベースの検索を行い料理レシピを出力する。伝達された情報が料理メニュー以外であった場合、プログラムを用いてPC操作を行う命令をPCに入力することで、実際にPC操作を行う。

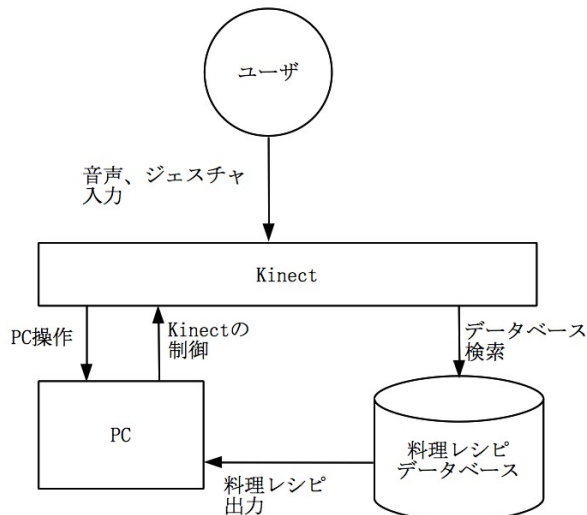


図1. 全体の構成

6.3 プログラム構成

本研究で作成したプログラム全体の流れは、図2のようになる。

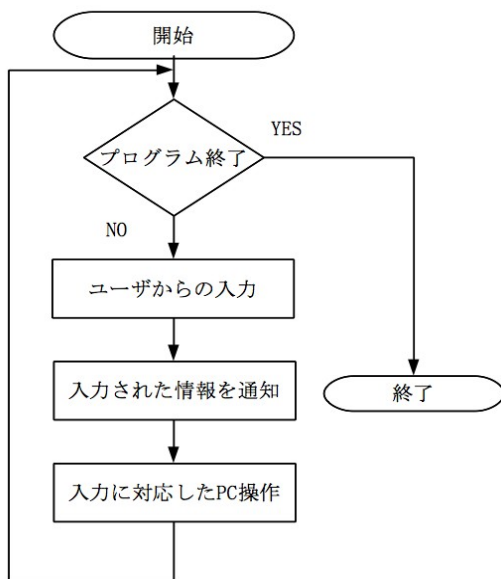


図2. プログラムの流れ

プログラムを起動すると終了操作が行われるか、ユーザからの何らかの入力があるまでは待機状態を維持し続ける。ユーザからの入力をKinectが受け取ったとき、ユーザからどのような入力を受け取ったかをユーザに通知し、受信した情報を元にPC内での制御を随時行う。この動作は、ユーザがプログラムを終了するまで、情報の送受信とPCの制御を繰り返し行う。ユーザから入力される情報はジェスチャと音声入力に分けられ、さらにジェスチャに関しては、右手と左手では入力で

きる操作を分けた。また、ユーザから料理のレシピが音声で入力されるとユーザが自分用に編集した料理データベースからユーザが発した料理レシピと一致したものを出力する。

7. プログラムの実行

料理をする上で最低限の機能を実装し、実際に料理を行った。図3に起動初期画面を示す。



図3. 実行画面

図3に表示されているA-Dは次の通りである。

- A. 右手をトレースしている赤い点
- B. 左手をトレースしている青い点
- C. ユーザが行ったジェスチャを表示するテキストボックス
- D. ユーザが発声した単語を表示するテキストボックス

A, Bは、左右の手をリアルタイムで追跡する青と赤の点を表示している。

C, Dのユーザが行ったジェスチャとユーザが発声した単語を表示するテキストボックスはユーザがPCを入力を行うたびに入力された事柄に対応したテキストボックスが変化する。

実際に料理を行ってみたいところ、鍋を火にかける時間などを測る為にタイマー機能を実装することでさらに使いやすいアプリケーションになると感じた。

8. おわりに

本研究ではKinectの機能を利用した料理者支援アプリケーションの提供を目的とし、アプリケーションの開発を行った。今回、事前検証を行ったことで、前述した学習機能とデータベース機能の実装を行うことの必要性を感じた。また、実際に料理を作るために最低限必要な機能を実装して料理を試みたところ、タイマー機能といった料理をする際に便利な機能を実装することも必要だと感じた。今後の予定としては、上記の機能の実装について検討していくとともに、料理をする際にどんな機能があれば便利であるかのアンケートをとり、その機能を順次実装していく予定である。

参考文献

- [1] 中村薫ほか、Kinect for Windows SDK プログラミングC#編