

音源情報と音声認識を用いたシリアルリンクの制御

永田研究室 F112039 孫 明洋

1. 目的

現在世界中では様々なロボットが使われている。知能型ロボットはある程度自律的に行動し、人間と仲良く同じ環境で作業する必要がある為に周りの情報や空間を認識することが重要である。身体的ハンディキャップがある人でも簡単に利用できるインタフェースを実現するためには、音声情報や音源情報の利用も有効である。本研究で取り扱う情報は Kinect センサで測定出来る音源情報と音声情報である。6 自由度の機構を有するシリアルリンクを製作し、音源方向と音声情報の認識機能を用いて音声により制御できるシステムを試作し、評価した。

2. 研究内容

小型の位置指令型サーボモータを関節に使用し、それらをリンクパーツで接続することで 6 自由度を有するシリアルリンクを製作した。設計した制御システムと試作したシステムをそれぞれ、図 1 と図 2 に示す。PC からシリアル通信でサーボコントローラに目標の関節角度を送信することで各関節をフィードフォワード制御できる。Kinect には 4 つのマイクが内蔵されており、それらを併用することで音が発生した方向を検知できる。C 言語で記述したプログラム内では Kinect for Windows ソフトウェア開発キット(SDK)を用いたが、USE_AUDIO 定数を定義してから NtKinect.h をインクルードすると NtKinect.lib 内の音声関係の関数や変数を有効に出来る。NtKinect.lib に含まれる setAudio() という関数を呼び出すと、推定された音源の方向が変数 beamAngle にセットされる。beamAngle には音声の方向(左右への角度)が格納され、正面を 0 度として、左側をマイナス、右側をプラスとする $-50 < \theta < 50$ 度の範囲の値をとる。Kinect で測定した音声データをパソコンに転送し、音声認識を開始する。音声認識エンジンにはマイクロソフト社の SpeechRecognitionEngine を用いた。認識方法には、自作した指定命令との一致による認識 (Recognition of command) と、自由形式の音声認識 (Recognition of free-form dictation) の 2 つのカテゴリがある。今回の研究では Recognition of command を用いた。命令を記述した辞書ファイルを作成し、イベントが発生したら認識結果を取り出せるようにした。実際の制御系では、音源データから得られた角度情報をモータ 0 に出力し、さらに音声データから認識された音声テキストに対応した動作命令をサーボモータ 1 から 5 の目標角度に変換し、出力させることで 6 つモータをそれぞれリアルタイムに制御できるようにした。

3. 結果

Kinect の音源検知機能と音声認識機能を利用して、操作者が発する音声に応じてシリアルリンクマニピュレータを動作させるシステムを試作し、動作実験を通して設計したソフトウェアの妥当性と有効性を確認した。今回は使用した音声認識エンジンの性能により、入力が英語の場合、十回の試行で三回程度出力が得られないか誤認識といった不具合が発生したため、今後はより高性能な音声認識エンジンを使用したい。

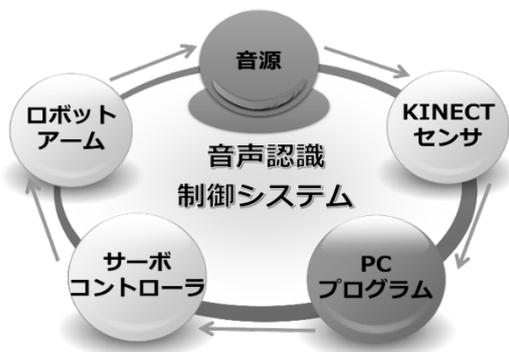


図 1 設計した制御システム

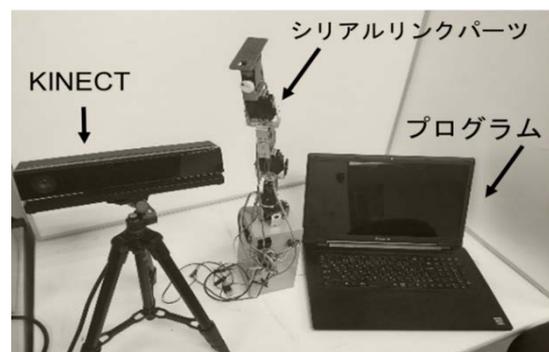


図 2 製作したシステム