

小型飛行ロボットによるオンボード音声認識 Onboard Sound Recognition using Small Aerial Robot

三好 賢聖
Kensho Miyoshi

堀 浩一
Koichi Hori

東京大学大学院工学系研究科航空宇宙工学専攻
Department of Aeronautics and Astronautics, The University of Tokyo

Recent and rapid advancement of robotic technology is bringing aerial robots closer to tasks and applications. Natural interaction of aerial robots is expected to be applied in wide areas from disaster countermeasures to entertainment. However there was only a method of using external microphone to realize audio-based interaction between human and aerial robots. We propose audio-based interaction using onboard microphone to realize closer and directer interaction.

1. はじめに

飛行ロボットと人間との直感的なインタラクション (Natural Interaction, NI) に関する研究が近年広く行われるようになり、災害対策からエンターテインメントまで幅広いシーンでの応用が期待されている。従来研究には飛行ロボットが空から捉える映像を用いたものが多くみられるが、本研究は飛行ロボットの「視覚」ではなく「聴覚」に関する研究である。

ユーザが携帯型マイクに音声を送り、音声認識により飛行ロボットを無線操縦する等といった音声インターフェイスについては従来に研究がなされてきた。しかし、飛行ロボット自体が音声センサを備えることによる周囲環境とのインタラクションの可能性についてはまだまだ未開拓であり、本研究はこの可能性を探索するものである。

次節以降、飛行ロボットとの NI について関連研究を紹介・考察し、オンボード音声センサによる音声認識、及びそれを用いたインタラクションについて実装したプロトタイプを交えて説明する。

2. 関連研究

飛行ロボットの NI に関する従来研究には、ジェスチャーなどの画像ベースのアプローチと、声などの音声ベースのアプローチという主となる2つの手法がある。更に、それらの画像及び音声センサ部がオンボードであるか否かという2通りに分類がなされる。HTA(Heavier-than-air)型飛行ロボットに関して、その従来研究を紹介する。

Monajjemi らはクアドコプターとそのオンボードカメラを用いてジェスチャーによるインタラクションを実現した[Monajjemi 2013]。Ng らはクアドコプターと設置型カメラを用いてジェスチャーによるインタラクションを実現した[Ng 2011]。前者の手法は、設置型カメラに比べ飛行ロボットのペイロード等のスペックに制約を受ける一方、飛行ロボットはその位置に関わらず周囲環境と画像ベースのインタラクションが可能である。後者の手法では、画像取得及び画像処理が飛行ロボットと分離されたシステムで行われるため、飛行ロボットのスペックに制約を受けることがない。しかしながら、ユーザは設置されたカメラの前でのみ飛行ロボットに対しアクションを行うことが許されており、人間にとってのインタラクションの可能な領域がカメラの視野に限定される。

Quigley らは固定翼の小型飛行ロボットに対し、PDA(携帯情報端末)を用いた音声入力による遠隔操作を行った[Quigley

2004]。PDA は地上で利用され、PDA によって得られた音声はコマンドに変換され、コマンドは飛行ロボットへ無線送信される。この手法はユーザが飛行ロボットを遠隔操縦する上で有効であるが、一方で飛行ロボット自体には音声センサは装着されていないため、飛行ロボットは自身の周囲環境の音声に反応することはできない。

以上のように HTA 型飛行ロボットの NI において、飛行ロボット自体が音声センサを搭載し、その周囲環境とインタラクションを行といったアプローチは殆どみられない。

3. オンボードマイクによる音声認識

我々が実現したのは、飛行ロボットのオンボードマイクを用いた音声認識及び、これを用いた周囲環境とのインタラクションである。設置型あるいは携帯型マイク等のデバイスを介することがないため、特定の操縦者を選ぶことがない。エンターテインメント等の分野では、複数人が同時に参加できる形のインタラクションが可能になる。災害救助等の利用シーンにおいては、周囲の画像情報に加え音声情報の収集に用いられることが考えられる。



図 1. 手のひらサイズのクアドコプター

4. 実装

本節では、我々の実装したプロトタイプを紹介する。飛行ロボットとして、此村らの手のひらサイズクアドコプター[Konomura 2013, 2014]を用いた(図 1)。本機はモータから対角のモータまでが 12cm, 重量は 70g で、ソフトウェアシステムは Linux のプラットフォーム上に構築されている。マイクはデジタルシリコンマイクロフォン(図 2)を用いた。

連絡先: 三好賢聖, 東京大学大学院工学系研究科航空宇宙工学専攻 堀研究室. miyoshikensho@ailab.t.u-tokyo.ac.jp

4.1 離陸時

離陸時には、飛行ロボット自体のノイズ以外に音声入力を阻害するものがない。離陸時の音声認識部には、オープンソースの汎用大語彙連続音声認識エンジン Julius[Lee 2009]を採用した。Julius はリアルタイムの音声認識処理を実行することができ、発音辞書や言語モデル等を自由に組み替えることができる。

本研究において、ユーザ辞書はコマンド群および、飛行ロボットの呼称の単語で構成されており、1)呼称のみ、2)コマンドのみ、3)呼称+コマンドという三種類の文法を認識する。実際にコマンドを実行するのは 3)の音声入力を得られたときのみとしている。これは周囲の騒音の誤検知の対策である。



図 2. オンボードマイク

4.2 飛行中

飛行中には、クアドコプター自体のプロペラによる騒音によるノイズが存在する。この騒音を周波数解析すると、図 3 のように人間の声の周波数帯に近い 100Hz～1kHz あたりに大きく分布していることがわかる。一方、3kHz 付近の周波数帯にはノイズがあまり見られない。そこで、3kHz 程度の音が出せるホイッスルを用いた実験を行った。音声入力から FIR フィルタ(図 5)を用いてバンドパスフィルタをかけ、窓関数にはカイザー窓関数を用いた。結果、プロペラ音による大きなノイズがある中、ホイッスル音を聞き分けることに成功した。ホイッスル音と着陸コマンドに対応させることにより、飛行中の飛行ロボットにオンボードマイクを用いた着陸指示が可能になった。

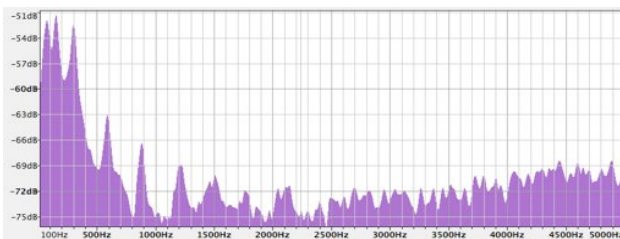


図 3. プロペラ音(スペクトラム表示)

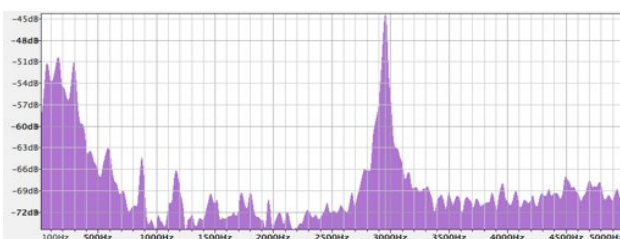


図 4. プロペラ音とホイッスル音(スペクトラム表示)

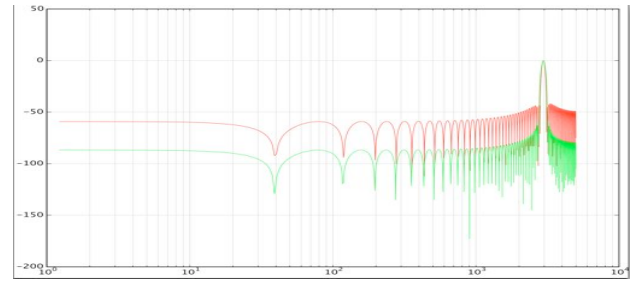


図 5. FIR フィルタ(緑:窓関数をかけたもの)

5. 結論

オンボードマイク及び MCU による音声認識による飛行ロボットとのインタラクションを実現した。エンターテインメントやナビゲーションなどの分野で、飛行ロボットと人間のより近く直接的なインタラクションを可能にすることが考えられる。また、情報収集の観点からは、従来盛んに研究が行われている空撮(画像情報)に加え、周波数帯の制限はあるものの音声情報の収集にも活用される可能性が見出されたといえる。

また現在は 3kHz 帯のホイッスル音が着陸コマンドのみに対応されているが、音の長さやその組み合わせパターンを聞き分けることで多種のコマンドに対応させ、加えてプロペラ音の存在下で人の声を抽出することにも取り組んでいく予定である。

参考文献

- [Lee 2009] Lee, A., Kawahara, T.: Recent development of open-source speech recognition engine Julius, APSIPA, pp. 131-137, 2009.
- [Konomura 2013] Konomura, R., Hori, K.: Designing Hardware and Software Systems Toward Very Compact and Fully Autonomous Quadrotors, IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2013.
- [Konomura 2014] Konomura, R., Hori, K.: Visual 3D Self Localization with 8 Gram Circuit Board for Very Compact and Fully Autonomous Unmanned Aerial Vehicles, 2014 IEEE International Conference on Robotics and Automation, 2014 (to appear)
- [Monajjemi 2013] Monajjemi, V., et al: HRI in the sky: Creating and commanding teams of UAVs with a Vision-mediated Gestural Interface, Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on IEEE, pp. 617-623, 2013
- [Ng 2011] Ng, W., et al: Collocated interaction with flying robots, RO-MAN/IEEE pp. 143-149, 2011
- [Quigley 2004] Quigley, M., et al: Semi-autonomous human-UAV interfaces for fixed-wing mini-UAVs, IEEE/RSJ International Conference on IEEE, pp. 2457-2462, 2004