

複数ドローンとロボットの協調による群アクティブ聴覚システム

Swarm Active Audition System with Robots and Drones

中臺 一博^{1*} 公文 誠² 佐々木 洋子³ 干場 功太郎¹ Yen Benjamin¹

¹ 東京科学大学/ Institute of Science Tokyo

² 熊本大学/ Kumamoto University

³ 産業技術総合研究所/ AIST

Abstract: 本稿では、複数のドローンとロボットが連携して、要救助者に由来する音源を検出することにより、要救助者を発見する群アクティブ聴覚システム (Swarm Active Audition System with Drones and Robots (SAAS-RD)) を提案する。地震などの広域災害現場における搜索救助活動の初期段階では、現場状況の把握を通じた搜索者の安全確保が第一優先課題であり、そうした状況下で広範囲を効率的に搜索することは容易ではない。本システムは、こうしたニーズにこたえるために提案した。まず、SAAS-RD のコンセプトとシステムアーキテクチャを紹介し、実機による性能評価実験、およびケーススタディを通じて、その有効性を実証する。さらに、災害対応に従事する搜索救助者へのインタビューを通じて、ユーザ視点から、SAAS-RD に求められる要件や必要機能を明らかにするとともに、現状の技術との潜在的なギャップ分析を行う。

1 はじめに

災害救助において、生存者の搜索は最も重要なタスクの一つであり、災害発生から 72 時間以内に救助できなかった場合、生存確率が著しく低下する「72 時間の壁」が知られている。被災地では交通網が寸断され、緊急車両も容易に通れない状況にあるため、ドローンによる搜索が有効な手段となる。ドローンにカメラを搭載し、カメラの情報をもとに搜索を行うことは有効な手段であるが、がれきに埋もれていたり、霧などで視界が悪い状況では、その有効性は限定的である。

このような状況を踏まえ、ロボット聴覚 [1] の延長として、マイクロホンアレイを搭載したドローンが遭難者の探索を行う「ドローン聴覚」が提案された。奥谷ら [2] は、ロボット聴覚用に開発された一般化特異値分解ベースの MUSIC 法 (GSVD-MUSIC) [3] を導入して、ドローン聴覚タスクにおける有効性を示した。また、Basiri らは、方位角・仰角の音源定位が可能な 4ch マイクロホンアレイを提案、実際に UAV (Unmanned Aerial Vehicle) に搭載してその有効性を示している [4]。大畑らは、GSVD-MUSIC 法にドローン雑音の逐次推定を組み込んだ iGSVD-MUSIC を提案した [5]。さらに、古川らは、ロータの回転情報からガウス過程でロータ雑音を逐次推定する手法を提案し、収録音を用いたシミュレーションでその有効性を示した [6]。

2014 年から始まった内閣府 ImPACT(革新的研究開発推進プログラム)の「タフ・ロボティクス・チャレンジ (TRC)」では、ドローン聴覚は災害地の要救助者発見に重要な技術であることから、特に音源定位と音源追跡の分野において広範な研究が行われ、ドローン聴覚分野は著しい進展を見せた。基本的なオフライン処理技術が確立すると、実環境下でも頑健でオンラインでも動作する技術の開発に注目が集まり、マイクロホンアレイの設計 [7]、3D 音源定位 [8]、および複数音源対応 [9] といった研究が行われ、最終的にオンラインでのデモへと発展した [10, 11]。この間、国外では、音源定位や追跡にフォーカスしたドローン聴覚研究の報告はほとんどなく、Wang らによる時間周波数マスクを使用した音源定位 [12] や Lauzon らによる地上設置のマイクを使用したドローンの位置推定研究 [13] など数例を数えるのみであった。この理由の一つは、DREGON データセット [14] のようなドローン聴覚用のデータセットが存在しなかったことが挙げられる。

一方で、ドローン聴覚用の音源分離と音声強調は、信号処理に基づく手法 [15, 16, 17, 18, 19, 20] から、深層学習に基づく手法 [21] に至るまで広く研究されている。国内では、ImPACT TRC を中心に、音源定位・追跡に続くステップとして、音源分離・識別との統合研究が行われている [22, 23, 24]。これらの研究で提案された手法の多くは、当時、大きな進展が見られた深層学習を用いた手法であり、信号処理と組み合わせた手法 [22, 24]、深層学習のみに基づく手法 [23] が報告されている。

*連絡先：東京科学大学
〒152-8552 東京都目黒区大岡山 2-12-1
nakadai@ra.sc.e.titech.ac.jp

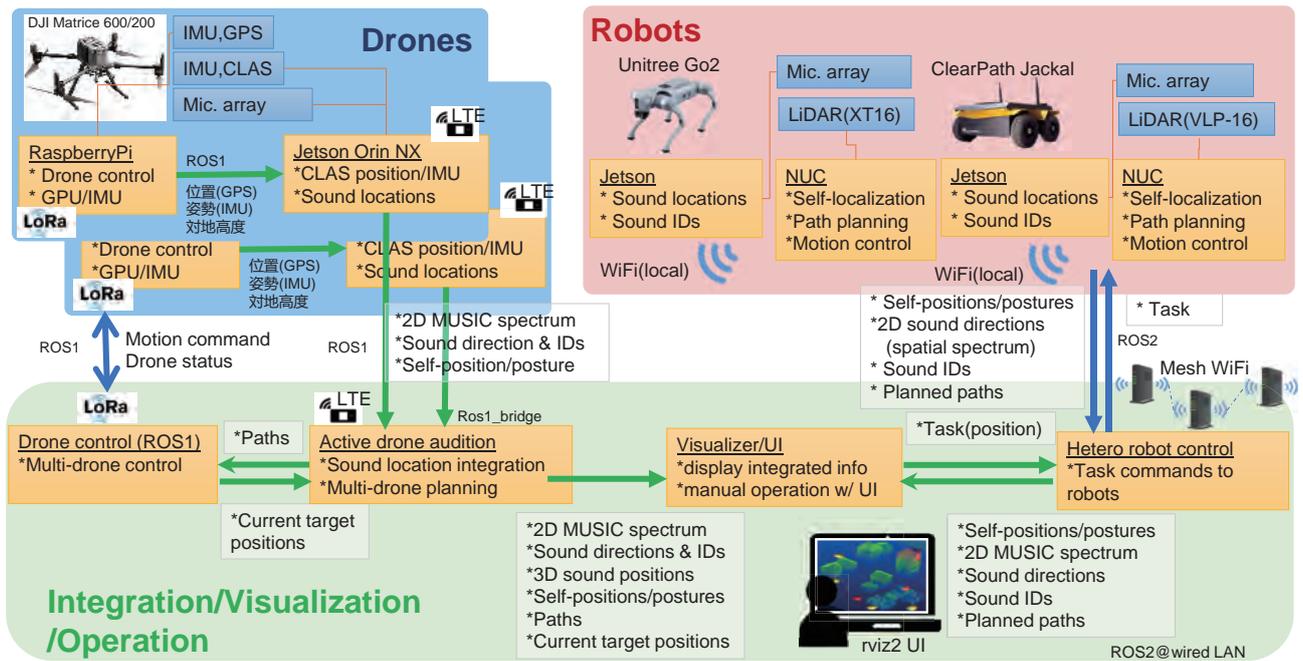


図 1: 群アクティブ聴覚システムのアーキテクチャ

近年、屋内使用を目的に音声認識の実現を試みる野心的な取り組みが行われている [25]。また、ロボット聴覚で提案されている、動作雑音の影響を低減するため、いったん止まって耳を澄ますという「stop-perceive-act」戦略 [1] の可能性を探る研究も行われている。例えば、ローターを停止させたまま滑空できるグライダーを利用して、ローターが停止中の静かな期間に音源探索を行うという、ユニークで実用的なアプローチが報告されている [26]。

三次元的に音源定位・追跡を行う手法も報告されている [10, 11, 9, 26]。しかし、この方法は2つの強い仮定を置いている。1つは、音源が地上にあるという仮定 [10, 11] であり、もう1つは、音源定位のために複数回の観測を行うという仮定である [9, 26]。一つ目の仮定は、救助が必要な人がビルの高層階にいる可能性があることを考えると、要救助者捜索には適さない。二つ目の仮定は、要救助者を発見するまでに要する時間が長くなる、搜索範囲が狭くなってしまふ、といった制約があり、現実的ではない。

これらの課題に対処するために、複数のドローンを同時に使用する研究が行われている [27, 28, 29]。複数のドローンを使用することで、前述の仮定を置くことなく、三角測量法を適用することで 3D 位置推定が可能となる。山田らは、この方法の限界として、複数の観測から得られる音源方向は必ずしも一点に交わらないという問題があることを指摘し、ガウスフィルタを使用してこの問題を解決する手法を提案した [27]。この手

法は、各ドローンに搭載されたマイクロホンアレイから得られる MUSIC スペクトルを統合して、パーティクルフィルタで追跡を行う PAFIM に発展した [28]。さらに、前処理として、ウィナーフィルタを導入して雑音耐性を向上するアプローチも報告されている [29]。

山田らは、こうした一連の研究をさらに一歩進めて、同時に、複数のドローンが能動的に最適な位置取りを行い、目的音源の定位精度を向上させる「アクティブドローン聴覚」の概念を提案した [28]。一方で、アクティブドローン聴覚の有効性評価に関しては、録音された音響信号を合成して入力とするオフラインの数値シミュレーションに限定されていた。

以上を踏まえ、本稿では、複数台のドローンに基づく音源定位・追跡を拡張し、複数のドローンおよびロボット群を使用する群アクティブ聴覚システム (Swarm Active Audition System with Robots and Drones (SAAS-RD)) を提案する。さらに、こうしたシステムが実際の災害シナリオで潜在的に有用であるかを確認するため、ヒアリングを通じた調査を実施した。本稿の貢献は、以下のとおりである。

- 災害現場における要救助者の捜索を目的とした SAAS-RD の提案
- 構築した SAAS-RD の性能評価
- ユーザ視点からの要求仕様、および必要機能の明確化と現状とのギャップ分析



図 2: DJI Matrice 600 Pro(左) と DJI Matrice 210 RTK V(右). 各ポール先端に、軽量特殊ウレタン製風防を付けた 16ch 球形マイク・アレイ (白色) を設置



図 3: Clearpath JACKAL(左) と Unitree Go2(右). 16ch の円筒形マイクアレイと LiDAR を搭載.

2 群アクティブ聴覚システム

図 1 に、SAAS-RD のアーキテクチャを示す。主に **ドローン群**、**ロボット群**、**統合/可視化/操作 (Integration, Visualization, and Operation, IVO)** の 3 つのブロックから構成されている。

ドローン群には、図 2 に示す 2 機のドローン DJI Matrice 600 Pro (6 枚プロペラ), DJI Matrice 210 RTK V (4 枚プロペラ) を用いている。複数のドローンを同時に使用することで 3D 音源定位を可能にし、探索の高速化と広域化を実現する。各ドローンには、16ch マイクロホンアレイ (図 2 参照) と、処理プロセッサとして Jetson Orin NX が搭載されており、Jetson 上で MUSIC スペクトルの算出を行う。MUSIC スペクトルは、水平角 72 方向、仰角 19 方向の計 1,368 方向のステアリングベクトルに対して 500ms ごとに計算され、その結果が、SAAS-RD の基地局側のシステムに実装されているアクティブドローン聴覚モジュールに送られる。このモジュールは複数のドローンから得られる MUSIC スペクトルを統合し、音源位置推定、音源追跡、およびドローンの最適配置計画を行う。最適配置計画は、LoRA を通して各ドローンに搭載されている Raspberry Pi に送信され、Raspberry Pi から、ドローン制御が行われる。

ロボット群には、図 3 に示す 2 機の異種ロボット Clearpath JACKAL (4 輪), Unitree Go2 (4 脚) を用いる。これらのロボットが、オペレータからの探索指示を契機に探索を開始する。まず、ドローン群の探索結果から、要救助者がいる可能性が高い場所をオペレータ

が割り出し、その位置を操作画面を通じて SAAS-RD に指定する。すると、SAAS-RD は、事前構築した 3D 地図と指定位置から、その場所を探索するのに最も適したロボットを各ロボットの走破性能や特徴をもとに決定し、そのロボットに対して目的位置を送信する。次に、目的位置を受信したロボットは、目的位置にたどり着くための経路をグラフ探索ベースの動作計画に基づいて搭載されている NUC を用いて作成し [30]、ロボットに動作指令を行う。ロボットは動作中、搭載 LiDAR を用いて 3D 地図を実時間で更新しながら [31]、更新された地図に対して、都度グラフ探索に基づく経路計画を行う [30] ことができるため、完全自律で目的地まで移動することが可能である。各ロボットには WEB カメラも搭載されており、カメラを通じてオペレータが要救助者を発見することで、要救助者発見タスクを完遂するものとする。また、ドローンと同様にマイクロホンアレイが搭載されており、音源定位や識別機能も有しているが、本稿ではその機能の統合には至っていない。今後の拡張機能である。

IVO は、システムのロバスト性とユーザの利便性を保証する。システムのロバスト性については、低遅延でロバストなデータ通信を実現するために 3 つの対策を講じる。第一に、複数の無線通信経路の組み合わせを採用する。具体的には、**ドローン群**と **IVO** 間の通信には、WiFi, LTE, LoRa など複数の通信経路を併用し、通信と接続のロバスト性を確保する。また、**ロボット群**と **IVO** 間の通信を円滑にするため、Mesh WiFi を採用し、広域かつ透過的な通信を実現する。さらに、**IVO** 内では、安定した通信を行うために、有線接続を採用している。

第二に、各通信経路に固有の帯域幅の制限を考慮したデータ圧縮を行う。例えば、**IVO** の表示は、人間のオペレータが大まかな状況を把握できればよいだけなので、正確な情報を伝える必要はなく、ユーザの理解を妨げない程度に情報量を減らしてもよい。また、情報の離散化も必ずしも均等に行う必要はない。情報量の劣化が問題にならない範囲では、離散化の刻みを大きくし、情報量の劣化が問題になる範囲では細かく離散化することで、知的なデータ圧縮を行っている。

最後に、データ配信サービス (DDS) を活用することで、大規模システムにも安定的に対応できることから、ROS2 を用いる。ただし、現状では、ドローンの組み込みシステムの制約から、**ドローン群内**と**ドローン群**と **IVO** 間の通信に ROS1 を使用している。

ユーザインタフェース (UI) を設計については、ユーザの利便性を高めるため、システムに不慣れなユーザが一目で状況を把握できるような設計としている。具体的には、ドローン雑音など不要な音源は表示しないもしくは目立たない色で表示し、音声など重要と思われる音源は、3D 地図上に音源種類ごとに色分けし、信

頼度を考慮して径の異なる楕円体として表示する機能を実装している。これにより、多くの音源候補から重要な候補が短時間で認識できる UI となっている。ただし、現時点ではドローン群、ロボット群の音源識別機能を利用していないため、外れ値的に表れる音源位置をドローン雑音として非表示にする以外は、同色での表示となっている。人間のオペレーターは、集中的に検索すべきエリアの指定を GUI を通じて行うなど、必要に応じたシステムへの介入が常に可能な設計となっている。

3 システム性能評価

システムの性能について、以下の点で評価を行った。実験 1-4 は、主にドローン群を用いたドローン聴覚に関する性能実験であり、実験 5 が統合システムの評価実験に相当する。

- 実験 1 目的音源の音量に対する検出性能
- 実験 2 環境雑音に対する頑健性
- 実験 3 遮蔽物存在下の検出性能
- 実験 4 被覆範囲や探索速度に関する性能
- 実験 5 実機によるケーススタディ (統合システム試験)

評価はいずれも福島県いわき市に位置する「ながとイノベーションセンター」にて実機で行った。同センターは、廃校になった小学校をそのまま利用した施設であるため、主に校庭でドローンを飛行させて評価を行った。音源は、主に学生 2 名の音声を用いた。ただし、遮蔽物存在下での検出性能については、レスキュー用の笛、金属をたたく打撃音も利用した。評価指標は音源検出率であり、出力音が正しく検出できた場合を成功、そうでない場合を失敗として、成功した回数を試行回数で割った値を音源検出率とした。各音源は、3 回ずつ出力を行った。なお、同センターは国道に面しているため、暗騒音は比較的大きく 50-55dBA であった。また、2 台のドローンがホバリングしている場合の暗騒音は 70dBA 程度であった。

実験 1 では、距離 20m を隔てて、高度 8m でホバリングしている 2 台のドローンを用いた音源検出実験を行った。音源と 2 台のドローンの中点位置までの距離を 10-40m まで 10m おきに、音源音量を 70-95dBA まで 5dBA おきに変えて実験した。音量は、話者の 1m 先に設置した騒音計で計測しながら目標音量になるよう発声してもらった。また、ドローン群にも Stop-Perceive-Act 戦略を導入することを考慮して、2 台のドローンを地面に着地させた静止状態で、同条件で実験を行った。ただし、音源音量は、55dBA-95dBA に範囲を広げて実験を行った。

実験 2 の周囲の騒音には、隣接する国道を走るトラック等の走行音や風雑音が含まれる環境雑音の収録音を用いた。実際にドローンで収録したシーンにこの収録雑音を率を変えて加算することで、70dBA~85dBA まで暗騒音を変化させた実験を行った。実験 1 と同様に高度 8m でホバリングさせた場合と地面に着地させた場合の 2 通りの実験を行い、音源の音量は約 100dBA、ホバリング時のドローン間の距離は 20m、各ドローンの中間地点と音源の距離は 10m、着陸時のドローン間の距離は 10m、各ドローンの中間地点と音源の距離は 5m とした。なお、この実験では収録音を用いてはいるが、オフラインシミュレーションで評価を行った。

実験 3 は、遮蔽物に対する性能である。距離 20m を隔てて、高度 18m でホバリングしている 2 台のドローンで 2F の音源を検出するタスクを実機で行った。ドローンが 2 階の軒の庇に隠れてしまう場合、音源とドローンの間に大きな樹木がある場合、2F の窓のカーテンを閉めた場合、窓を少しだけ開いた状態にした場合、完全に窓を閉めた場合を遮蔽条件として選択した。音声の音量は 85dBA と 95dBA、レスキュー用の笛は純音に近く 95dBA 程度の音量、金属の打撃音は 90dBA 程度の音量であった。

実験 4 は、被覆範囲に関する実験である。実験 1 の結果から、一回の探索で 40m (音源までの距離) × 20m (ドローン間距離) 程度の被覆があることが判明したので、移動しながら、これを繰り返し行うことで、75m×75m の範囲をどの程度の時間で被覆できるかを音源検出率と共に実機で測定した。

実験 5 は、統合テストである。ケーススタディとして、校庭に 1 名、校舎の 2F に 1 名の計 2 名の要救助者がいることを想定した要救助者発見タスクを実施した。タスクは、この 2 名の音声をドローンで検出し、オペレーターにその場所を呈示し、さらにオペレーターの指示に従ってロボットの WEB カメラで 2 名の姿を補足することをもってタスク成功と定義した。

3.1 実験結果

図 4 に実験 1 の結果を示す。図 4a) から、ホバリング中では、75dBA (10m) - 90dBA (40m) の音量が必要であることがわかる。ホバリング中の環境雑音が 70dBA 程度であるので、+5 - +15dBA の音源であれば検出が可能であることになる。音量的には大きい声に分類されるため、実シナリオを考えると、要救助者が大きな声を出せる状況にある、災害初期で比較的元気な状態である場合に有効な技術であるといえる。そのような状況下であれば、一回の探索で 20m (ドローン間距離) × 40m (音源までの距離) の範囲を被覆可能といえる。これは 50 坪の家 5 軒分に相当する広さで

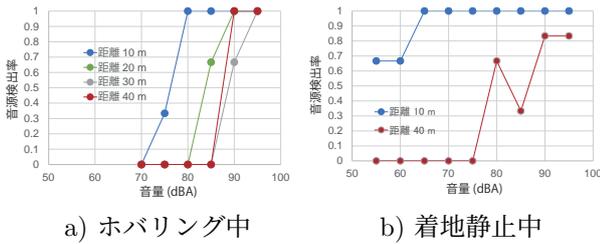


図 4: 音量に対する音源検出率

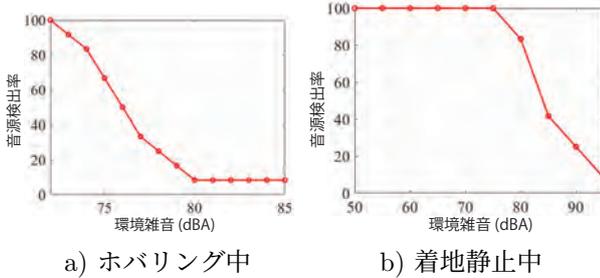


図 5: 環境雑音に対する音源検出率

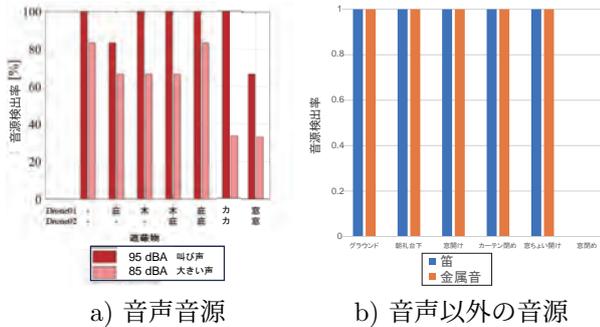


図 6: 遮蔽に対する頑健性

ある。一方で、ドローンを停止させて耳を澄ませるといふ戦略を併せて考えると、図 4b) からわかるように、音源が 10m 以内であれば 55 dBa 程度の小さな声まで拾うことが可能であることがわかる。つまり、このような戦略を併せてとることで本技術の適用の幅が大きく広がると考えられる。

図 5 に実験 2 の結果を示す。音源検出率 50% のところを見ると、ホバリング中で 76 dBa、停止中で 83 dBa 程度の環境騒音までは音源の検出が可能であるといえる。騒音規制法では建設工事現場は、85 dBa 以下の環境であることが求められていることを考えると、その範囲はほぼカバーできていると考えられる。ただし、実際の災害現場の騒音環境はこれ以上に過酷である可能性が高く、さらなる性能の向上が必要であるといえる。

図 5 に実験 3 の結果を示す。図 5a) は音声音源の場合の結果であるが、窓を閉めたような完全遮蔽された場合であっても、大きい声を出せる状況であれば検出可能であることがわかる。一方、図 5b) に示す笛や金属打撃音では、窓閉めのような完全遮蔽ケースでは、

表 1: 75 m × 75 m の被覆タスクの性能

平均音源検出率	探索時間
86.4%	3:06

音源の発見は難しい。一方で少しでも窓が開いていれば、音が漏れ出ることができ、検出が可能であることがわかる。音源による違いは、周波数特性の違いによるものと推測されるが、詳細を知るにはさらなる調査が必要である。

表 1 に実験 4 の結果を示す。実験 1 より、要救助者が大きな声を出せる好条件下では、20 m × 40 m 程度の被覆があることが分かったので、これを 7 回繰り返すことで 75 m × 75 m に相当する範囲の探索に対する性能を計測した。ドローンは約 2 m/s で飛行させたが、40 m 進むごとにホバリングさせて、音源の探索を 10 秒程度行うというタスクを行った。また、ホバリングだけではなく、飛行中も音源探索は継続して行った。結果として、平均音源検出率で 86.4% と高い検出率を得ることができた。また、被覆にかかった時間は 3 分程度であり、迅速に広範囲を探索するための有効な技術であることがわかる。

図 7 に実験 5 のシーンのスナップショットを示す。この実験は SAAS-RD のシステム統合テストの位置づけで、ライブデモ形式で行った。2 台のドローンによる音源検出だけでなく、オペレータが定位された音源情報をもとに、その音源を詳しく調査する命令を GUI を用いて簡単に行うことができること、その命令を受けた SAAS-RD が指定位置を調査するために最適なロボットを自動的に選択して、完全自律でその位置まで移動できること、指定位置に到着したロボットが搭載されている WEB カメラで要救助者に見立てた人を補足し、それをオペレータが認知できることを示すことができた。また、システムの通信遅延も、このセッティングでは平均 1 秒程度であることがわかり、システムとしての動作、有効性を示すことができた。

4 災害現場ドローン作業経験者へのヒアリング

ドローン聴覚は、過酷な環境下での音響信号処理と機械学習技術の有効性への関心から、シーズ主導で研究がおこなわれてきた。これは、ドローン聴覚の目的が要救助者の探索であることに端を発しているが、実際にドローン聴覚技術や SAAS-RD が、このような状況において本当に有用なのか、必要機能を網羅しているのかといったことをユーザ視点で検証することはできていない。本節では、これらの検証を試みる。

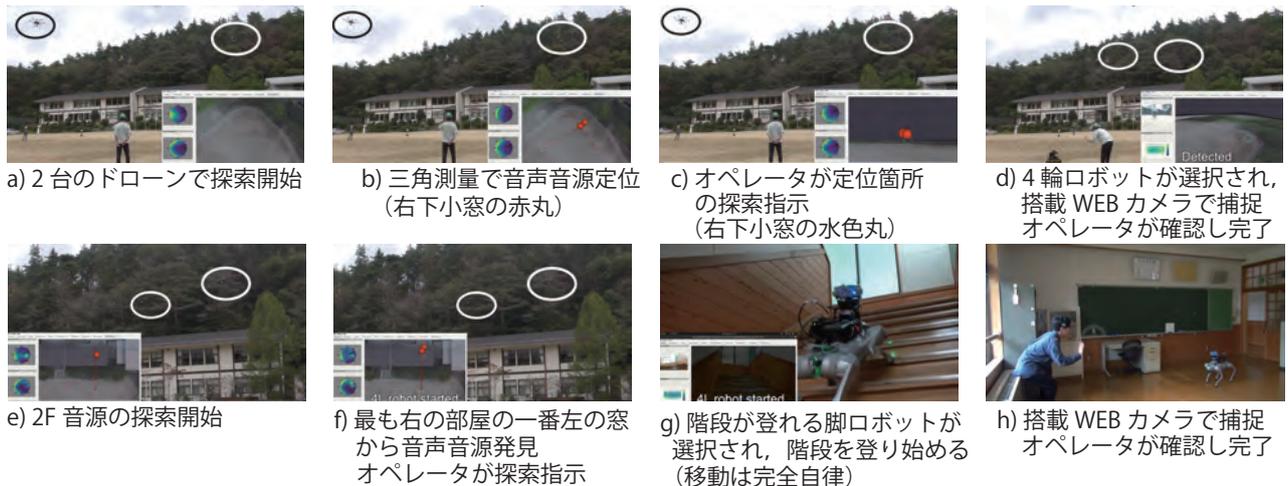


図 7: ケーススタディ (2 音源の発見)

表 2: ヒアリング項目 (抜粋)

種類	内容
概要	災害救助でドローンを利用する目的は？ ドローン聴覚が貢献しそうなケースとは？ 現場でのドローン利用状況は？
データ	だれが、いつ、どのように使用するか？ カメラとサーモセンサの課題は？ 聴覚技術を過去に利用した経験は？
環境	現場周辺状況は？ データ通信の課題は？ 現場には他にどんな音源があるか？
ユース ケース	平常時のドローンの利用法は？ 災害救助以外でドローン聴覚技術 が使える可能性は？

このため、これまで災害現場でのドローン活用実績や、災害現場でのドローンへの期待などをヒアリングし、ドローン聴覚技術への期待や他の技術との棲み分けに対する現場の考え方を調査した。ヒアリング対象は、2024年1月に発生した能登半島地震の救助活動に携わった経験を持つ県消防本部2と、ドローン運用会社2の計4者である。表2は、災害現場の課題、災害現場の環境、ロボット聴覚技術へのニーズ、他技術との比較などを考慮し、選定したインタビュー項目を列挙したものである。インタビューは1者あたり1時間程度で行い、単なる Yes/No 質問だけではなく、具体的な内容を総合的に聴取した。

災害現場の課題 現場で要救助者搜索の経験を実際に持っていたのは1者だけであった。理由として、ほとんどの救助要請が災害発生から2週間以上たってから

であり、災害発生後3日間という要救助者搜索のタイミングでの現場に入ることができなかったことが挙げられる。ドローンは状況把握のためのオルソ画像作成のために使われることが多く、それがドローンの有効な活用例であるというコンセンサスがあることが背景にあるようである。また、要救助者搜索の際の最優先事項は、搜索側のメンバの安全であり、二次災害を防ぐため、安全を確保してから搜索を行うため、時間がかかってしまう。

災害現場の環境 環境条件は場所によって大きく異なるが、典型的な状況について以下に示す。

- オルソ画像生成では、ドローンは通常、飛行時間10~20分、速度2~3m/s、高度50~100mで運用する。
- 現場にはさまざまな音源があり、うるさい状況である。主な騒音源は風、人、重機、ヘリコプターなどである。
- 火災の場合は、炎や消火用の水しぶき音がドローン騒音をかき消すほど大きい。

ロボット聴覚技術の必要性 ポジティブなコメント：

- 倒壊家屋からの救助は、最終的には人の聴覚を頼りに搜索する。人は安全を確保してからでないと搜索を開始できないため、ドローン聴覚で補うことができれば有効。
- スピーカーを使って呼びかけ、聞き分け技術をもっと積極的に搜索に使うべき。
- 災害時にドローンで医薬品などの物資輸送する際に相手が確実に受け取ったかどうかを確認する場合など、音声で相互コミュニケーションができるとう有効。

- 山岳救助現場では、木々が遮蔽物となりカメラや赤外でも発見が難しいケースがあり、要救助者の声がドローンに届けば有用。この場合にも音声による相互コミュニケーションができることさらに有効。

否定的な意見としては、要救助者は、衰弱した状態であることも多く、大きい声を出すことは難しい。通常の会話よりも小さい声を検出できるのであれば有効であるが、ドローン騒音下ではむしろかしいのではないかといった声が寄せられた。

他の技術との比較 現場では、聴覚をはじめとして人の感覚に頼ることが多い。ただし、有効性が期待できる技術を導入することには積極的であり、画像探査装置や電磁波探査装置は要救助者発見に利用している。しかし、これらの装置は長所だけではなく、検索範囲が制限されるといった短所もあり、現時点では、最適なソリューションを提供する技術はない。

5 考察

本節では、調査結果の考察、SAAS-RD が提供する機能がユーザにとって不可欠なものなのか、あるいはSAAS-RD に現在欠けている機能があるのかを分析する。

5.1 ヒアリング結果の考察

上述したように示唆に富んだ情報が得られた。現状、要救助者捜索は人手に頼っており、ドローンを用いた要救助者捜索の有効性を認識しつつも、ドローンの利用が盛んに行われているわけでないことが現場の課題である。この問題は、災害救助の構造的な問題や捜索者側の意識による部分も大きい。また捜索のために利用可能なセンシング装置もまだ未成熟であることから、要救助捜索のための技術を早急に確立して、展開を図ることが大事である。このためには平時から利用可能なタスクで技術に慣れ親しんでもらうことが重要であると考えられる。

災害現場環境では、要救助者由来の音源以外にも、様々な音源が存在することが重要な知見として得られた。これは、音源定位・追跡に加え、音源分離・識別が重要な技術になりうることを示している。オルソ画像と比較すると、ドローン聴覚による要救助者捜索は、低い高度での実施が必要である。低空飛行で安全を確保するためには、自動制御システム、動的マッピング、動的経路計画など、ドローン操縦者を支援するための技術の強化も不可欠である。

災害シナリオにおけるドローン聴覚技術の有用性については、現場従事者のコンセンサスが得られた。また、ドローンと要救助者間のコミュニケーションを機能への要望が強いことが分かった。また、要救助者の声が小さい場合、ドローン聴覚が有効に機能しないことを懸念する声も得られた。

5.2 SAAS-RD 機能に関する考察

前述の議論を踏まえ、現行のSAAS-RD が災害現場に従事するユーザにとって必要な機能を備えているか、また不足している機能はないかを検証する。

まず、ドローン聴覚機能を検証する。**ドローン群**の主な機能は、3D 音源定位、高速・広域探索である。最初の2つの機能は、捜索救助の現場は状況が変化しやすく、救助者の位置が事前に把握できないこと、人手による捜索活動に依存している現状では、捜索そのものよりも、捜索者の安全が優先されること、を考慮すれば、現場で本格的に利用された際には、災害初期の捜索を担う有用な技術になることが明らかである。小音量の音声を検出する必要性を考えると、将来的には、**ドローン群**にも「stop-perceive-act」戦略を導入する必要があるといえる。現時点では、実装されていないが、音源識別機能は、現場に多数の音源が存在することを考えると、効率的な捜索には不可欠であるといえる。

ロボット群は、異種ロボットを用いた高い走破性、経路計画による動的マップ生成機能を有する。前者は、災害現場の状況に事前に何ら仮定が置けない場合に、また、後者は、災害現場の具体的な状況がすぐには分からない場合に有効な機能である。**ロボット群**の音響処理は現時点では、システムに統合されいないため、有効に使われてはいないが、**ドローン群**と比べると低速で移動し、「stop-perceive-act」戦略がとりやすいため、高い定位解像度が期待できる。一方で、懸念点としては、**ドローン群**では検知できない小さな音量の音を**ロボット群**でどれだけ検知できるかが挙げられる。この検証には、さらなる評価が必要だろう。

IVO は、複数の無線通信チャネルの併用、通信帯域幅の制約を考慮したデータ圧縮技術、ROS2の採用により、ロバスト性を高めるように設計している。これらの対策の実施を阻害するような意見は確認されなかった。災害現場の作業者は、基本的に工学技術の専門家ではない。そのため、一目で状況を把握できるようなUIの提供が不可欠である。レスキュー隊と要救助者のドローンを介しての相互音声通信については強い要望があった。今後のUI実装で検討すべき項目である。

6 おわりに

本稿では、複数のドローンとロボットが協調して要救助者を検索する群アクティブ聴覚システム (SAAS-RD) を提案した。その要素機能として、複数ドローンによる音源定位に着目してその性能を明らかにするとともに、SAAS-RD システムとしての有効性もケーススタディを通じて実証した。さらに、実際の災害現場で SAAS-RD が有用かどうかを検証するため、災害現場での本研究に従事した関係者にヒアリングを行った。その結果として、SAAS-RD の機能は基本的にユーザーの要望に沿ったものであることがわかった。小音量の声を検知する場合を懸念する意見も得られたが、これは、ロボット群を活用するなど、SAAS-RD 全体で解決できるものであると考えている。また、要望が強かった音声通話機能については、今後導入を検討していきたい。

謝辞

本研究は、JSPS 科研費 JP19KK0260, JP20H00475 および JP23K11160 の助成を受けた。また、M-KUBOS に関する部分については、科学技術振興機構戦略的研究推進事業 (JST), CREST, JPMJCR19K1 の支援を受けた。

参考文献

- [1] K. Nakadai, et al. Active audition for humanoid. In *AAAI*, pp. 832–839. AAAI, 2000.
- [2] K. Okutani, et al. Outdoor auditory scene analysis using a moving microphone array embedded in a quadcopter. In *IROS*, pp. 3288–3293, 2012.
- [3] K. Nakamura, et al. Real-time super-resolution sound source localization for robots. In *IROS*, pp. 694–699. IEEE, 2012.
- [4] M. Basiri, et al. Robust acoustic source localization of emergency signals from micro air vehicles. In *IEEE/RSJ IROS*, pp. 4737–4742, 2012.
- [5] T. Ohata, et al. Improvement in outdoor sound source detection using a quadrotor-embedded microphone array. In *IROS*, pp. 1902–1907, 2014.
- [6] K. Furukawa, et al. Noise correlation matrix estimation for improving sound source localization by multirotor UAV. In *IROS*, pp. 3943–3948, 2013.
- [7] T. Ishiki and M. Kumon. Design model of microphone arrays for multirotor helicopters. In *IROS*, pp. 6143–6148, 2015.
- [8] K. Washizaki, et al. Position estimation of sound source on ground by multirotor helicopter with microphone array. In *IROS*, pp. 1980–1985, 2016.
- [9] K. Yamada, et al. Belief-driven control policy of a drone with microphones for multiple sound source search. In *IROS*, pp. 5326–5332, 2019.
- [10] K. Nakadai, et al. Development, deployment and applications of robot audition open source software HARK. *JRM*, Vol. 29, No. 1, pp. 16–25, 2017.
- [11] K. Hoshiba, et al. Design and assessment of sound source localization system with a UAV-embedded microphone array. *JRM*, Vol. 29, No. 1, pp. 154–167, 2017.
- [12] L. Wang and A. Cavallaro. Time-frequency processing for sound source localization from a micro aerial vehicle. In *ICASSP*, pp. 496–500, 2017.
- [13] J.-S. Lauzon, et al. Localization of RW-UAVs using particle filtering over distributed microphone arrays. In *IROS*, pp. 2479–2484, 2017.
- [14] M. Strauss, et al. DREGON: Dataset and Methods for UAV-Embedded Sound Source Localization. In *IROS*, pp. 5735–5742. IEEE, 2018.
- [15] Y. Hioka, et al. Speech enhancement using a microphone array mounted on an unmanned aerial vehicle. In *IWAENC*, pp. 1–5. IEEE, 2016.
- [16] S. Yoon, et al. Two-stage adaptive noise reduction system for broadcasting multicopters. In *2016 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 219–222, 2016.
- [17] L. Wang and A. Cavallaro. Acoustic sensing from a multirotor drone. *IEEE Sensors Journal*, Vol. 18, No. 11, pp. 4570–4582, 2018.
- [18] Z.-W. Tan, et al. An efficient dilated convolutional neural network for UAV noise reduction at low input SNR. In *APSIPA ASC*, pp. 1885–1892, 2019.
- [19] B. Yen, et al. Multi-sensory sound source enhancement for unmanned aerial vehicle recordings. *Applied Acoustics*, Vol. 189, p. 108590, 2022.
- [20] W. N. Manamperi, et al. Drone audition: Audio signal enhancement from drone embedded microphones using multichannel Wiener filtering and Gaussian-mixture based post-filtering. *Applied Acoustics*, Vol. 216, p. 109818, 2024.
- [21] L. Wang and A. Cavallaro. Deep learning assisted time-frequency processing for speech enhancement on drones. *IEEE Trans. Emerg. Topics Comput. Intell.*, Vol. 5, No. 6, pp. 871–881, 2021.
- [22] S. Uemura, et al. Outdoor acoustic event identification using sound source separation and deep learning with a quadrotor-embedded microphone array. In *ICAM*, 2015.
- [23] T. Morito, et al. Partially shared deep neural network in sound source separation and identification using a UAV-embedded microphone array. In *IROS*, pp. 1299–1304, 2016.
- [24] O. Sugiyama, et al. Outdoor acoustic event identification with DNN using a quadrotor-embedded microphone array. *JRM*, Vol. 29, No. 1, pp. 188–197, 2017.
- [25] T. Fuhrman, et al. An interactive indoor drone assistant. In *IROS*, pp. 6052–6057, 2019.
- [26] M. Kumon, et al. Alternating drive-and-glide flight navigation of a kiteplane for sound source position estimation. In *IROS*, pp. 2114–2120, 2021.
- [27] T. Yamada, et al. Assessment of sound source tracking using multiple drones equipped with multiple microphone arrays. *Int. J. Environ. Res. Public Health*, Vol. 18, No. 17, 2021.
- [28] T. Yamada, et al. Placement planning for sound source tracking in active drone audition. *Drones*, Vol. 7, No. 7, 2023.
- [29] B. Yen, et al. A performance assessment on rotor noise-informed active multidrone sound source tracking methods. *Drones*, Vol. 8, No. 6, 2024.
- [30] Y. Sato, et al. STP4: spatio temporal path planning based on pedestrian trajectory prediction in dense crowds. In *RtTA*, 2022.
- [31] S. Nijima, et al. Semi-automatic town-scale 3D mapping using building information from publicly available maps. *IEEE Access*, Vol. 10, pp. 32244–32254, 2022.