Audio-based Swarming Micro Aerial Vehicles

acceptée sur proposition du jury:

Prof Alcherio Martinoli, président du jury
Prof Dario Floreano, directeur de thèse
Prof Pedro Lima, directeur de thèse
Prof Colin Jones, rapporteur
Prof Stefano Stramigioli, rapporteur
Prof Patric Jensfelt, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

TÉCNICO
LISBOA

Lausanne, EPFL, 2014

Acknowledgements

Firstly, I would like to thank my two supervisors, Professor Dario Floreano and Professor Pedro Lima for their support and guidance during my PhD studies. I am grateful to them for providing me with a great research environment and for teaching me to think out of the box. I thank Dario for raising my interest in the domain of aerial robotics, for encouraging me to pursue acoustics as a potential solution to my research questions and for teaching me how to best present my work. I am very grateful to Pedro for all his valuable feedbacks and comments, for his encouragements and for giving me the confidence in the times where I needed it the most.

Next, I would like to thank my friend, supervisor and colleague Doctor Felix Schill for all his great expert advices, for all the fruitful discussions and for all his insights and eagerness to help. It was an honour to work alongside someone with a high level of technical and scientific expertise in the field. I am very thankful for all his help and support and for all the things I learned from him. This work would not have been possible without hist support.

Next, I would like to express my gratitude to the members of my jury, Professor Patrick Jensfelt, Professor Stefano Stramigioli and Professor Colin Jones for taking the time and effort to read my thesis, and for their valuable feedback in improving the final manuscript. I am grateful to Professor Alcherio Martinoli for presiding the jury and also for all the valuable information I learned form him in the course of distributed intelligent systems that contributed to the development of this work. Special thanks to Patrick who introduced and raised my interest in the field of robotics and for all his supervisions and support during my master studies that encouraged me to pursue this path.

I am grateful to both EPFL and IST for providing a great studying and working environment, to the Foundation for Science and Technology (FCT) and Laboratory of Intelligent Systems (LIS) for funding this research. I want to thank Grégoire Heitz and Yannick Gasser for all their assistance and help in building some of the robotic platforms used in this work. I am thankful to Sabine Hauert, Adrien Briod, Severin Leven, Nicolas Dousse and Maja Varga for their assistance, fruitful discussions and insights. And I would like to thank all my colleagues and friends from the Laboratory of Intelligent Systems in Lausanne and the Institute for Systems and Robotics in Lisbon. Special thanks to my two office mates Dr Andrea Maesani and Trevis Alleyne for all their encouragements in difficult times, and to Dr Stefano Mintchev and Dr Ilya

Abstract

Employing a swarm of independently controlled Micro Aerial Vehicles (MAVs), instead of a single flying robot, can increase the robustness and efficiency of many aerial coverage missions through parallelism, redundancy and cooperation. Swarms of autonomous MAVs also show great potential in many diverse applications. In a search and rescue mission, a group of MAVs could quickly reach disaster areas by flying over obstacles, cluttered and inaccessible terrains, and work in parallel to detect and locate people that are in need of help. However, swarms of MAVs have so far mostly been demonstrated in simulation or in few real works in well prepared environments and with the aid of external systems that are either impractical or not always available. This is due to the multiple challenges that still remain in the design of truly autonomous MAV teams, where on-board solutions that are independent of any external systems and that can satisfy the strict constraints imposed on these small, lightweight, inexpensive and safe robots are required. This thesis contributes by proposing solutions to some of these challenges to assist with the future deployment of MAV swarms for real missions, and in particular for search and rescue operations.

Designing a group of autonomous MAVs requires addressing challenges such as self local-ization and relative positioning. Individual's knowledge about their three dimensional position is required for allowing MAVs to navigate to different points in space and to make decisions based on their positions. Furthermore, robots within an aerial swarm need to interact with each other and to work together towards achievement of a global goal, where knowledge about the position of other swarm members is necessary for overcoming challenges such as, moving in formations, avoiding inter-robot collisions, and achieving distributed search and coverage tasks. Due to the strict constraints imposed on the MAVs in terms of size, weight, 3D coverage, processing power, power consumption and price, there are not many technological possibilities that could provide individuals with the self-localization and relative positioning information without the aid of external systems. Inspired by natural swarms, where complex cooperative behaviours are acquired from local actions of individuals that rely entirely on their local senses, and furthermore, inspired from the sense of hearing among many animal groups which use sound for localization purposes, we propose an on-board audio-based system for allowing individuals in an MAV swarm to use sound waves for obtaining the relative positioning, and furthermore, the self-localization information. We show that not only such a system fully satisfies the constraints of MAVs and is capable of obtaining these information, but also provides additional important opportunities, such as the detection and localization of crucial acoustic targets in the environment. Operating during night time, through foliage

Abstract

and in adverse weather conditions such as fog, dust and smoke, and detection and collision avoidance with non-cooperative noise-emitting aerial platforms are some of the potential advantages of audio-based swarming MAVs.

In this thesis, we firstly describe an on-board sound-source localization system and novel methods for the innovative idea of localizing emergency sound sources, or other narrowband sound sources in general, on the ground from airborne MAVs. For MAVs involved in a search and rescue mission, the ability to locate the source of distress sound signals, such as the sound of an emergency whistle blown by a person in need of help or the sound of a personal alarm, is significantly important and would allow fast localization of victims and rescuers during night time and in fog, dust, smoke, dense forests and otherwise cluttered environments. We propose multiple methods for overcoming the ambiguities related with localizing narrowband sound sources.

Furthermore, we present the on-board audio-based relative positioning system for providing individuals in an MAV swarm with information about the position of other MAVs in their vicinity. We initially describe a passive method that exploits only the engine sound of other robots, in the absence of the self-engine noise, to measure their relative directions. We then extend this method to overcome some of its limitations, by proposing active acoustic signalling where individuals generate a chirping sound similar to the sound of birds to assist others in obtaining their positions. A method based on fractional Fourier transform (FrFT) is used by robots to identify and extract sounds of simultaneous chirping robots in the neighbourhood. We then describe an estimator based on particle filters that fuses the relative bearing measurements with information about the motion of the robots, provided by their onboard sensors, to also obtain an estimate about the relative range of the robots.

Finally, we present a cooperative method to address the self-localization problem for a team of MAVs, while accommodating the motion constraints of flying robots, where individuals obtain their three dimensional positions through perceiving a sound-emitting beacon MAV. All methods are based on coherence testing among signals of a small on-board microphone array, to measure the probable direction of incoming sound waves, and estimators for robust estimation of the desired information throughout time. In addition, all solutions are independent of any external system and rely entirely on the swarm itself.

Overall, this thesis presents innovative methods for contributing to the three challenges of target localization, relative positioning and self localization in a swarm of micro air vehicles using sound as the main source of information.

Key words: Aerial robots, micro aerial vehicles, multi robots, swarm, relative positioning system, localization, search and rescue, sound-source localization, acoustic targets, microphone array, narrowband sounds, emergency sounds, audio-based system, MAV, UAV.

Résumé

Le fait d'utiliser une flotte de robots volant, chacun contrôlé de manière indépendante, permet d'augmenter de manière considérable la robustesse et l'efficacité de nombreuses missions de couverture aérienne. Ceci est dû au parallélisme des taches effectuées, à la redondance et à leur coopération. Les flottes de robots volants, autonomes, démontrent également leurs forts potentiels à travers des applications diverses et variées. Dans les missions de secourismes, une telle flotte pourrait rapidement atteindre une zone sinistrée, en survolant les obstacles au sol et les milieux denses voir inaccessibles par les voies terrestres. Cette flotte pourrait alors se partager la zone de recherche, afin de détecter et localiser plus rapidement de potentielles victimes nécessitant une assistance. Cependant, jusque-là, l'utilisation de ces flottes de robots volants n'ont démontré leurs potentielles, qu'à travers des simulations ou des environnements réels, très contraints, reposant sur des interventions humaines ou des technologies qui ne sont pas toujours applicables à de telles situations. Ceci est dû au challenge propre à la construction de robots vraiment autonomes, ou les systèmes embarqués doivent être indépendants de tous systèmes extérieurs et satisfaire aux contraintes de poids, taille, prix et sureté requis pour de telles missions. Cette thèse vient donc contribuée, à ce sujet, en proposant des solutions à certains des challenges évoqués précédemment, afin d'assister le futur déploiement de telles flottes dans des missions réelles, particulièrement pour celles de secourisme.

Le design d'une flotte de robots volant requiert de résoudre certains challenges tels que l'auto-localisation et le positionnement relatif aux autres. La connaissance, par chaque robot volant, de sa position en trois dimensions, est nécessaire afin de pouvoir naviguer dans l'espace. De plus, chaque robot volant a besoin d'interagir et collaborer avec les autres. Et ce, en vue de réaliser certaines tâches ou la connaissance des positions des autres individus est nécessaire, comme voler en formation ou d'assurer la non-collision entre robot, ou encore d'explorer ou couvrir une zone de manière distribuée. Les possibilités technologiques disponibles, de nos jours, permettant l'auto-localisation et le positionnement relatif sont limitées. Ceci est principalement dû aux strictes contraintes en termes de taille, poids, couverture en trois dimensions, capacité des processeurs embarqués, consommation énergétique et prix, qui s'appliquent aux robots volants, Dans la nature, des collaborations complexes ont pu être observées, ou l'action locale de chaque individu repose entièrement sur leurs capacités sensorielles locales. A ce titre, l'ouïe est utilisée chez de nombreux animaux pour se localiser les uns les autres. S'inspirant de ces deux derniers points, nous proposons une solution basée sur l'utilisation des ondes sonores, permettant à chaque robot volant de se positionner locale-ment par rapport aux autres afin de s'auto-localiser. Nous ne nous limitons pas à démontrer,

ici, qu'une telle solution répond aux contraintes liées à l'utilisation de robots volants et est capable d'obtenir ces informations. Elle offre également de nombreuses autres opportunités, telles que la détection et localisation de cibles acoustiques présentes dans cet environnement. Cette solution permet entre autre d'opérer de nuit, à travers des feuillages denses et dans des conditions météorologiques difficiles, telle que le brouillard, les nuages de poussières ou de fumées. Elle permet également de détecter et d'éviter des plateformes volantes non coopératives, grâce au son que celles-ci émettent naturellement.

Dans cette thèse, nous commençons par décrire une méthode embarquée de localisation d'une source sonore ainsi que la méthode innovatrice pour localiser, depuis un robot volant, un signal de détresse émis depuis le sol. Pour un robot volant impliqué dans une mission de secourisme, la possibilité de localiser un signal de détresse, tel que celui émis par un sifflet ou une alarme personnelle, est d'une grande importance. Elle permettrait ainsi de trouver rapidement des victimes ou sauveteurs et ceci quelques soient les conditions, même de nuit, par brouillard, malgré des nuages de poussières ou de fumée, à travers une forêt dense ou n'importe quel environnement compliqué. Nous proposons trois différentes méthodes pour résoudre le problème d'ambiguïté lié à la localisation de source sonore.

De plus, nous présentons la méthode embarquée permettant à un robot volant de se localiser par rapport aux autres en utilisant le son. Nous commençons par présenter une méthode passive pour déterminer la direction des autres robots, qui utilise seulement le bruit émis par leur moteur. Nous présentons ensuite une approche utilisant une émission active de son émis par chaque individu, comme le font les oiseaux pour se localiser les uns les autres. Une méthode basée sur la transformée de Fourrier fractionnaire est utilisée pour permettre d'identifier et discrétiser les sons émis simultanément par les robots voisins. Nous décrivons ensuite une estimation basée sur les filtres à particules qui mixe les informations d'orientations du son émis avec la trajectoire du robot, obtenu à partir de ses propres capteurs, pour également estimer la distance des robots voisins.

Enfin, nous présentons une méthode coopérative pour répondre aux problèmes d'auto-localisation pour une flotte de robots volants, ou ils ajustent leurs trajectoires selon la position en trois dimensions de la source sonore qu'ils perçoivent. Toutes ces méthodes s'appuient sur la cohérence des mesures prises par plusieurs microphones, afin d'en déduire la probable direction de l'onde sonore entrant et d'améliorer son estimation au cours du temps. Enfin, il nous semble important de rappeler que toutes les méthodes présentées sont indépendantes de systèmes extérieurs et ne reposent ainsi que sur la flotte de robots volants elle-même.

En conclusion, cette thèse présente une manière innovante de résoudre les trois problèmes suivants : localisation d'une cible, localisation relative et auto-localisation, d'une flotte de petits robots volants, utilisant le son comme information principale.

Mots clés : robots aériens, des micro véhicules aériens, robots multiples, essaim, système de positionnement relatif, la localisation, recherche et sauvetage, localisation de source sonore, cibles acoustiques, réseau de microphones, sons à spectres étroits, des sons d'urgence, le système basé sur l'audio, MAV, UAV .

Resumo

Usar um enxame de Micro Veículos Aéreos (MVAs) controlados independentemente, em vez de um único robô voador, pode aumentar a robustez e eficiência de várias missões de cobertura aérea, atrás de paralelismo, redundância e cooperação. Enxames de MVAs autónomos possuem também potencial em várias aplicações de diversos tipos. Numa missão de busca e salvamento, um grupo de MVAs pode chegar rapidamente a áreas de desastre, voando por cima de obstáculos, terrenos confuso e de difícil acesso, e trabalhar em paralelo para detectar e localizar pessoas necessitadas de auxílio. No entanto, enxames de MVAs têm até agora sido demonstrados principalmente em simulação ou em escassos trabalhos em ambientes reais, nos quais o ambiente de teste é bastante controlado e auxiliado por sistemas externos que são ou pouco práticos ou completamente indisponíveis em situações reais. Tal facto é devido aos múltiplos desafios que ainda permanecem na elaboração de equipas de MVAs verdadeiramente autónomas, onde soluções com sensores a bordo de MVAs que são independentes de quaisquer sistemas externos e que podem satisfazer as restrições rigorosas impostas sobre estes pequenos, leves, baratos e seguros robôs são requeridas. Esta tese contribui para este esforço, propondo soluções para alguns destes desafios, com o intuito de auxiliar a futura utilização de enxames de MVAs para missões reais, em especial para as operações de busca e salvamento.

Projectar um grupo de MVAs autónomos implica enfrentar desafios como auto-localização e posicionamento relativo. Conhecimento de cada indivíduo sobre a sua posição tridimensional é necessário para permitir que MVAs naveguem para diferentes pontos no espaço e tomem decisões com base nas suas posições. Para além disso, robôs operando dentro de um enxame aéreo precisam de interagir com outros membros do enxame e trabalhar em conjunto para a realização de um objectivo global, onde conhecimento sobre a posição dos outros membros do enxame é necessário para a superar desafios tais como navegação em formação, evitar colisões inter-robô e alcançar tarefas de busca e cobertura distribuídas. Devido às limitações rigorosas impostas aos MAVs em termos de tamanho, peso, cobertura 3D, poder de processamento, consumo de energia e preço, não há muitas possibilidades tecnológicas que possam fornecer os indivíduos auto-localização e informações de posicionamento relativo sem a ajuda de sistemas externos. Inspirado por enxames naturais, onde comportamentos cooperativos complexos são construídos a partir de acções locais de indivíduos que dependem inteiramente dos seus sensores locais, e no sentido da audição, usado em vários grupos de animais para fins de localização, propomos um sistema de sensores a bordo de MVAs baseado em áudio para permitir que membros de um enxame MAV possam utilizar ondas sonoras para

Resumo

a obtenção do posicionamento relativo e, além disso, informação de auto-localização. Mostramos que este sistema, não só satisfaz plenamente as limitações de MVAs e é capaz de obter estas informações, mas também oferece oportunidades adicionais importantes, tais como a detecção e localização de alvos acústicos cruciais no ambiente. O funcionamento durante o período nocturno, através de vegetação e em condições climáticas adversas, como neblina, poeira e fumo, e a detecção e prevenção de colisões com plataformas aéreas não-cooperativas e emissoras de ruído, são algumas das potenciais vantagens de enxames de MVAs baseados em áudio.

Nesta tese, descrevemos um sistema de localização baseado em sensores a bordo com fonte sonora e novos métodos para a ideia inovadora de localizar fontes de som de emergência no chão a partir de MVAs no ar. Para MVAs envolvidos em missões de busca e salvamento, a capacidade de localizar a fonte de sinais sonoros de emergência, como o som de um apito de emergência soprado por uma pessoa que precisa de ajuda ou o som de um alarme pessoal, é significativamente importante e permitiria a localização rápida de vítimas e socorristas durante a noite e na névoa, poeira, fumo, densas florestas e ambientes desordenados. Propomos três métodos diferentes para superar as ambiguidades relacionadas com a localização de fontes de som de emergência ou, em geral, fontes de som de banda estreita.

Seguidamente, apresentamos um sistema de posicionamento relativo baseado em sensores a bordo com fonte sonora para fornecer membros de um enxame de MVAs com informações sobre a posição de outros MVAs na sua vizinhança. Inicialmente, descrevemos um método passivo que explora apenas o som do motor de outros robôs, na ausência de ruído dos próprios motores, para medir as suas direcções relativas. De seguida estendemos esse método para superar algumas de suas limitações, propondo uma sinalização acústica activa onde os membros de um enxame MVAs geram um som semelhante ao chilrear de aves para ajudar os outros na obtenção de suas posições. Um método baseado em Transformada de Fourier Fraccionária (TFrF) é usado por robôs para identificar e extrair sons de chilrear de robôs simultâneos na vizinhança. Em seguida, descrevemos um estimador baseado em filtros de partículas que funde as medições de rolamento relativos, com informações sobre o movimento dos robôs, fornecidas pelos seus sensores a bordo, para obter também uma estimativa sobre a distância relativa dos robôs.

Por fim, apresentamos um método de cooperação para resolver o problema de auto-localização para uma equipa de MVAs, que considera as restrições de movimento de robôs voadores, onde os indivíduos obtêm suas três posições dimensionais através de perceber um MVA que funciona como farol emissor de som. Todos os métodos são baseados em testes de coerência entre os sinais de um pequeno conjunto de microfones a bordo, para medir a direcção provável de ondas sonoras de entrada, e estimadores para a estimativa robusta da informação desejada ao longo do tempo. Além disso, todas as soluções são independentes de qualquer sistema externo e dependem inteiramente do próprio enxame.

Palavras-chave: robôs aéreos, micro veículos aéreos, sistemas de múltiplos robôs, enxame, posicionamento relativo, localização, busca e salvamento, alvos acústicos, sons de banda estreita, localização baseada em som, sistema baseado em áudio, MVA, VANT.

Contents

Contents

List of Figures

1 Introduction

T HE goal of this thesis is to contribute to the field of aerial robotics by proposing solutions to some of the challenges faced in the design of autonomous swarms of Micro Air Vehicles (MAVs). We aim at proposing new methods to assist with the future deployment of MAV swarms for real missions and in particular for search and rescue operations. We propose novel methods that exploit sound as the main source of information in order to allow individual robots inside an aerial robotic swarm to obtain information about the position of themselves, information about the position of other swarm members, and to localize interesting acoustic targets in the environment. This chapter initially presents the background and motivation behind this work and describes some of the challenges associated with the design of autonomous MAV swarms. The state of the art is then reviewed describing the related work towards addressing the mentioned challenges. Finally an overview of the thesis is provided along with a description of the main contributions of this work.

Figure 1.1: Artistic view of audio based swarming MAVs

1.1 Background and motivation

Flying robots can rapidly reach areas of interest by flying over obstacles, cluttered and inaccessible terrains and are capable of providing elevated and bird's eye view sensing of target areas. Such important features make flying robots suitable for applications in many distinct science, defence and commercial environments. In search and rescue missions and in disaster situations similar to the Japan's 2001 earthquake and tsunami (see Figure 1.2), aerial robots could be employed to rapidly reach target areas in order to gather first critical information and to search for survivors until the arrival of ground units [Basiri et al., 2012]. Other applications of flying robots include environmental monitoring [Goktouan et al., 2010], aerial surveillance and mapping [Templeton et al., 2007], traffic monitoring [Puri, 2005] and infrastructure inspection [Sa and Corke, 2014, Katrasnik et al., 2008].

Employing a swarm of aerial robots for exploration and coverage missions, instead of a single flying robot, can increase the robustness and efficiency of such missions through redundancy, parallelism and cooperation. For example, in a search and rescue mission in a wide area incident, it is possible to obtain a rapid aerial coverage by dividing potential target areas among multiple flying robots working in parallel. In addition, sharing resources among multiple robots allows the development and employment of small, simple, inexpensive, robust and safe "Micro Air Vehicles" (MAVs) for many complex tasks that either require a large, expensive and complicated aerial platform or are beyond the ability of a single robot. Swarms of MAVs also introduce many interesting applications such a rapidly deployable communication network over a disaster area with damaged communication infrastructures (Figure 1.3) [Hauert et al., 2010]. Sensing and mapping of chemical clouds [Kovacina et al., 2002, Oyekan and Huosheng, 2009], searching for forest fires [Merino et al., 2006], aerial surveillance system [Beard et al., 2006], aerial transportation [Michael et al., 2011], building constructions [Lindsey et al., 2012] and detecting targets of interest [Ruini and Cangelosi, 2009, Sauter et al., 2005, Yang et al., 2005, Altshuler et al., 2008] are some of the other applications that are envisioned for MAV

Figure 1.2: Pictures of Tohoku earthquake and tsunami, illustrating the complexity of the environments for search and rescue operations in such disaster situations. Image Credits: Douglas Sprott, https://www.flickr.com/photos/dugspr
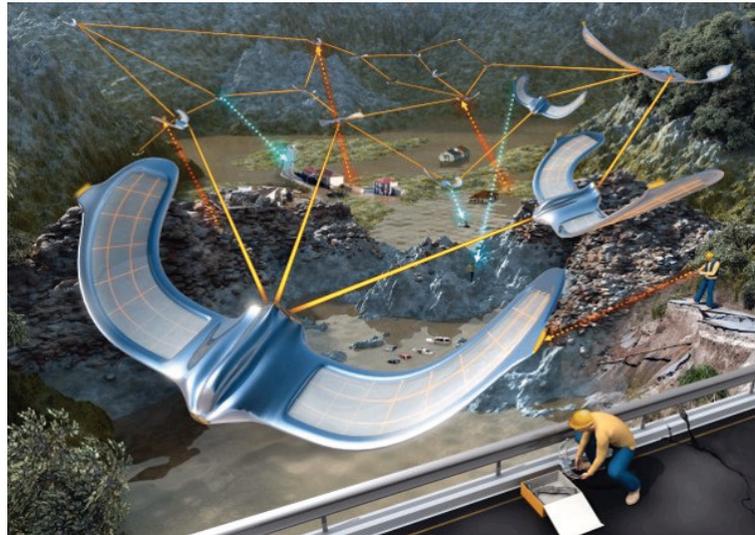
Figure 1.3: Swarming Micro Air Vehicles (SMAVs) could be employed to create a rapidly deployable wireless communication network over a disaster area to replace damaged communication infrastructures, allowing communication between different rescue teams and the mission coordinator and coordinating rescue efforts [Hauert et al., 2010]

swarms.

A rapid progress in the design and control of real multi MAV systems have been observed over the recent years [Kushleyev et al., 2013, Lindsey et al., 2012, Ritz et al., 2012, Hauert et al., 2011] that mostly exhibit autonomous operation of real swarms inside well prepared environments and with the aid of impractical external systems such as motion tracking cameras. Despite this progress, multi MAV systems have rarely been used so far for real mission scenarios. Two key challenges imposed in the design of MAV swarms that need to be considered before allowing their use for real missions are:

1. Self Localization

2. Relative Positioning

Self Localization is the problem of estimating the MAV's location relative to its environment. Individual's knowledge about their 3D position is essential for allowing MAVs to navigate autonomously to different points in space and to make decisions based on their current positions. Furthermore, robots within an aerial swarm are required to interact with each other and to work together towards achievement of a desired goal. This introduces additional problems such as inter robot collisions and formation control. A common idea that has been addressed throughout the literature of both the natural and artificial swarms is that individual's knowledge about the relative position of other swarm members is essential for achieving successful swarming behaviours [Reynolds, 1987, Pugh and Martinoli, 2006, Mataric, 1997]. For example, awareness about the relative range and/or bearing of neighbouring robots

Figure 1.4: Picture of a commercially available omnidirectional microphone sensor used in some part of this work, purchased at a price of around 1 Euro, that has a radius of 2mm and a weight of 0.04g

can allow a robot to maintain formations [Basiri et al., 2010, Moshtagh et al., 2009], and to decrease the risk of collisions [Carnie et al., 2006], with other team members.

A solution to the mentioned challenges for a swarm of MAVs must satisfy strict constraints imposed on the MAVs in terms of weight, size, power consumption, processing power, three-dimensional coverage and price. These hard constraints limit many successful solutions, that are available and are being used on ground robots or large aerial vehicles, to be transferred to MAVs. The most common approaches used for MAVs are based on Global Positioning System (GPS) sensors, to obtain the position of robots with the aid of GPS satellites, and a wireless communication network between robots, to share the obtained positions with each-other. However, GPS vulnerability is considered as one of the main problems that need to be solved before allowing MAVs to operate inside civilian airspace [James et al., 2001, Humphreys, 2012, Conte and Doherty, 2008]. GPS technologies are exposed to jamming and interferences [Pinker and Smith, 1999], have low update rate and resolution [Kernbach, 2013], and are impossible to use in indoors or cluttered environments where there is no direct line of sight with the GPS satellites [Siegwart and Nourbakhsh, 2004]. The lack of suitable solutions for self localization and relative positioning of MAV swarms that satisfies the MAV constraints and is independent of any external systems motivated this research study.

Inspired by natural swarms, where complex cooperative behaviours are acquired from local actions of individuals that rely entirely on their local senses [Mataric, 1993], and furthermore, inspired from the sense of hearing among many animal groups which use sound for localization purposes [Muller and Robert, 2001, Pollack, 2000, Farnsworth, 2005], we hypothesized that an on-board audio-based system is a promising solution that could potentially allow an MAV swarm to exploit sound waves for obtaining the relative positioning, self-localization and the autonomy required for real operations. Such a system will be based on low cost, small size, passive and omnidirectional microphone sensors, as shown in Figure 1.4, which clearly satisfy the mentioned MAV constraints.

We envision many potential advantages and applications for audio based MAV swarms. Firstly, they will be based on sound waves that are independent of illumination, weather conditions such as fog, dust and smoke, and can overcome obstacles throughout their ways, allowing

the operation of the swarm in night-time and through foliage. In addition, they could be employed to detect and locate many interesting acoustic targets in the environment. For example a team of MAVs could be employed in search and rescue missions to quickly detect and locate emergency acoustic sources in the area, such as the sound of a personal alarm or a safety whistle that is being blown by a person in need of help. Furthermore, audio-based MAVs could potentially detect, locate and avoid collision with other non cooperative noise-emitting aerial platforms since the engine of most flying platforms already generate sound while flying. Remote environmental noise monitoring, aerial surveillance and study of biological systems are some of the other applications we envision for MAV swarms with hearing capabilities. More potential applications are described later in Chapter 5.

To summarize, the lack of on-board solutions to relative positioning and localization, that is independent of any external systems and that can satisfy the constraints of MAVs, are some of the reasons limiting the employment of MAV swarms for real missions despite their great potential benefits. This motivate research on a new approach, where we hypothesized that an on-board audio-based system could be a promising solution for obtaining this information and a great asset for localizing targets from the air, such as locating victims in a rescue mission.

1.2 State of the art

This section briefly discusses the existing approaches to the problems of self localization in flying robots and relative positioning for multiple aerial robot systems along with a description of their advantages and limitations.

1.2.1 Localization methods

Robot localization is a well known problem faced in the design of autonomous mobile robots, that is described as the problem of determining the robots' positions relative to a single reference point. In general, robot localization methods that are addressed in the literature can be divided into two main categories:

1. Global Localization methods

2. Local Localization methods

In global localization methods, the position of robots relative to a single global reference frame is obtained with the assist of an external system. Using external colour vision cameras [Altuğ et al., 2005] or infrared 3D motion tracking cameras [Lupashin et al., 2010, Valenti et al., 2007] for indoor aerial robots, and using Global Positioning System (GPS) [Hauert et al., 2011] and wireless positioning beacons [Corke et al., 2005] for outdoor aerial robots, are some of the examples of global localization methods that are used in aerial robotics domain. The advantage with solutions based on this approach is the good accuracy that they usually provide

while having a low computational complexity. Their main drawback is the dependency on an external system that might not always be available in reality. Deployment of motion tracking cameras or wireless positioning beacons in the environment and in advance of each mission, if not impossible, is both costly and time-consuming. Furthermore, GPS sensors require a constant direct line of sight with the transmitting satellites and hence they can not be used indoors or in cluttered areas. GPS sensors are also vulnerable to interferences and jamming [Pinker and Smith, 1999], they have a poor vertical resolution of few tens of meters [Kernbach, 2013] and a low update rate of few updates per second. Robustness of aerial robots against GPS failures is considered as an essential feature that must be obtained before allowing the real operation of aerial robots inside civilian airspace [James et al., 2001].

Due to disadvantages of the first approach, much effort has been put into the design of local localization methods, where the position of robots are obtained locally, using onboard sensors, and independent of any external systems. In this group of methods, localization is achieved using probabilistic techniques and by only employing on-board proprioceptive and exteroceptive sensory information. The most common examples of this approach used on MAVs are the vision based SLAM (Simultaneous Localization and Mapping) algorithms that mainly use an onboard camera and an Inertial Measurement Unit(IMU) to map features in the environment and to localize the robot throughout time [Blosch et al., 2010, Bryson and Sukkarieh, 2007, Artieda et al., 2009, Weiss et al., 2011]. This group of methods are mostly used for indoor aerial robots where GPS signal are not available. The main drawback with the local localization methods is that they mainly require a high computational power and a high data storage for operation that is not always available, specially on small scale micro air vehicles. The need for real-time processing of high resolution and high frame-rate images, the dependency on illumination, visual contrast, weather conditions and the limited field of view of vision sensors and the errors caused due to the high or insufficient number of features in the images, the long displacement between loop closings and the fast dynamic nature of MAVs, are some of the major drawbacks of the visual SLAM methods for aerial robots [Weiss et al., 2011, Artieda et al., 2009].

1.2.2 Relative positioning

In natural swarms, e.g., flocking in birds and schooling in fishes, individuals mainly rely on information about the relative location of other swarm members for achieving swarming behaviours. Inspired from nature, several research works have shown that individual's knowledge of their local neighbours is sufficient for obtaining successful swarming of artificial agents [Reynolds, 1987, Pugh and Martinoli, 2006, Mataric, 1997, Burgard et al., 2000]. In the case of collective aerial systems, many spatial coordination methods are presented, mostly in simulation, that use relative positioning information to achieve tasks such as, formation flight [Basiri et al., 2010, Moshtagh et al., 2009, Anderson et al., 2008], mid air collision avoidance[Carnie et al., 2006, Shim et al., 2003] and optimal dispersion for sensor coverage [Parunak et al., 2003].

One way of obtaining the relative positioning information is to use a global localization system, described in Section 1.2.1, alongside a communication network [DeLima et al., 2006, Pack et al., 2006, Hauert et al., 2011]. This allows robots to obtain their positions with the assist of an external system and then share this information with each other to compute the relative positions. However, this method suffers from the dependency on an external system and other limitations of global localization methods explained in Section 1.2.1. In addition, the need for an effective communication network among all robots might result in high communication overhead that limit the scalability of solutions [Pugh et al., 2009] and introduces additional challenges such as routing and scheduling. Furthermore, this approach requires active cooperation of all individuals in communicating their locations, which could increase the chance of failures.

A more closer-to-nature solution to relative positioning is to directly measure the relative location of other robots using only on-board exteroceptive sensors. Solutions based on this approach are mostly developed for ground robots, that provide positions in two-dimensional space, and rely on sensors such as laser range finders [Simmons et al., 2000], infrared sensors [Pugh and Martinoli, 2006, Kemppainen et al., 2006, McLurkin and Smith, 2007] and cameras [Das et al., 2002]. However, since flying MAVs require a system with low cost, small size, low weight, low complexity and three-dimensional coverage, many of the successful systems implemented on ground robots can not be used on MAVs. Despite this, some effort has been done in transferring these solutions from ground robots to MAVS. Throughout the rest of this section we summarize the current state of the art in on-board sensor-based relative positioning systems developed for MAVs and investigate other existing sensor technologies as potential candidates for such a system.

Laser range scanners are one of the favourites in ground-based robotics, providing accurate planar scan of the robot's surrounding environment that are used to estimate and track the robot positions throughout time [Jensfelt and Christensen, 2001, Nüchter and Hertzberg, 2008, Brenneke et al., 2003]. However their bulkiness is considered as the main drawback which stops them from being used on MAVs. A 3D laser scanner with a limited field of view of 30x40 degrees have been used in [Scherer et al., 2007] on an outdoor helicopter with a payload of 29kgs. Mini laser range finders have been used by [Saunders et al., 2005, Kownacki, 2011], as shown in Figure 1.5.a, for detection of large static obstacles (trees and buildings) located directly in front of an MAV. These sensors provide accurate range measurements of obstacles located in front of the laser beam up to a few hundreds of meters away. A major drawback of such sensors is their single point/planar detection ability, which makes them a bad candidate for measuring the position of other MAVs in three-dimensional spaces. Also, these sensors are sensitive to ambient natural light and considered as power consuming sensors since they require emission of a laser beam for operation [Hebert, 2000].

Several works show the effectiveness of optical sensors for achieving fixed-obstacle and terrain avoidance in miniature aerial vehicles [Griffiths et al., 2006, Beyeler et al., 2009, Byrne et al., 2006]. Few works also investigate the use of cameras along with heavy and bulky hardware

(a)	(b)	(c)
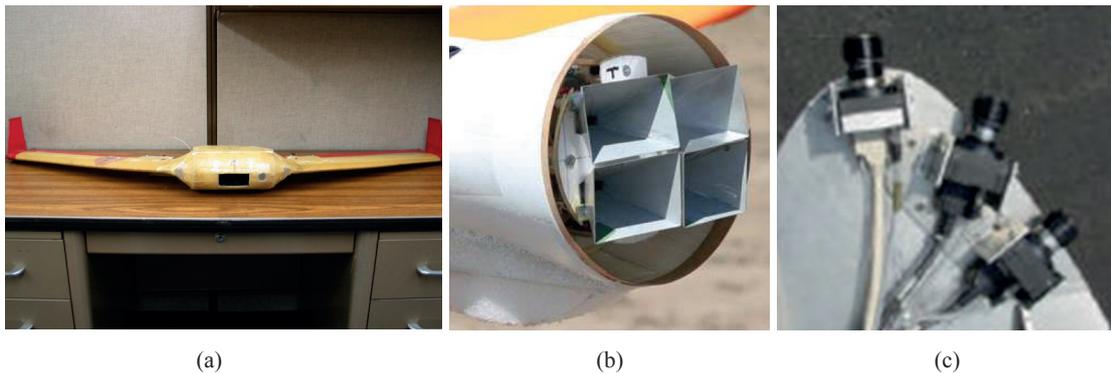
Figure 1.5: a) Mini laser range finder mounted in-front of an MAV for static obstacle detection [Saunders et al., 2005], b) Four Doppler radar transducers mounted in-front of an unmanned aerial vehicle to monitor the UAV's flight path for obstacles [Viquerat et al., 2008] , c) Three cameras mounted inside the nosecone of a small aircraft for detection of other aircraft [Utt et al., 2005]

on unmanned aerial vehicles (UAVs) for detecting the motion of other aircraft relative to the background scene and for computing their relative directions [Utt et al., 2005, Mejias et al., 2010]. Figure 1.5.c shows the optical cameras used in the work by [Utt et al., 2005] that is attached in-front of a small aircraft. Systems based on such sensors have a limited field of view and are highly dependent on light conditions and visual contrast. These systems also greatly suffer from missed or false detections when the target is located on non-uniform or cluttered backgrounds and also in the presence of vibrations and adverse weather conditions. Furthermore, image processing of high resolution and high frame rate images is considered a very computationally intensive task that require expensive and heavy hardware [Mejias et al., 2010] that are still challenging for use on MAVs.

Radar-based systems, such as the Traffic Collision Avoidance system (TCAS), are used for relative positioning between commercial aircrafts. Their operations are based on a transponder that sends out a radio message to the transponders of nearby aircrafts asking for their information. Once a reply message is received, it determines the range (from the time-of-flight of the message's round trip) and bearing (from directivity pattern of its directional antennas) of the corresponding aircraft. TCAS systems are not suitable for small scaled aerial platforms or MAVs as they are bulky, expensive and can only detect other transponder-equipped aircraft. More recently, advancements in radio communication modules with highly accurate, synchronized clocks has led to research in the time-of-flight range-only measurement of cooperative wireless sensor nodes [Lanzisera et al., 2006]. But so far, these modules have not been tested on flying platforms.

Small scale Doppler radar transducers, shown in Figure 1.5.b, are the basis of the sensor suite proposed in [Viquerat et al., 2008] for allowing an MAV to detect the presence and measure the relative bearing of colliding obstacles . The sensor suite has a weight of about 300 grams and

Figure 1.6: Sensor suit proposed by [Roberts et al., 2009] for an on-board relative positioning system for and indoor flying team of robots. This sensor suit weights $\approx$ 400 grams and consists of 160 IR transmitters and 48 photo diodes and is capable of providing accurate 3D range and bearing inside a range of 10 meters.

power consumption of 3.7 watts. However, having a small field of view of (30°), a low angular resolution of (15°) and a small operating range of 10 meters are some of the major drawbacks of this system.

Few recent works also show the use of ultrasound distance-measuring sensors on MAVs for measuring the altitude and/or avoiding collisions with large static obstacles in small indoor environments [Muller et al., 2014, Becker et al., 2012]. Similar to Doppler radar transducers, such sensors also consist of a transmitter and a receiver, and they operate by measuring the time difference between the time of transmission of a pulse of ultrasound wave and the time of receiving its echo reflected from a solid surface. These sensors are small and light-weight and suitable for using on MAVs. However, small working range ($< 10m$), small field of view (35°) and the dependency on the shape, density and material of detecting objects are some of the major disadvantages of ultrasound sensors, making them unsuitable for measuring the distance to other small scale micro air vehicles in three dimensional space.

To our best of knowledge, the only totally onboard relative positioning systems demonstrated for actual MAV teams are based on infra-red (IR) sensors [Melhuish et al., 2002, Roberts et al., 2009] and are used for indoor flying robots. The sensor suit proposed by [Roberts et al., 2009], illustrated in Figure 1.6, have been shown to provide accurate three dimensional relative range and bearing estimations, within a 10 meters range. This sensor suit consist of a total of 208 IR transmitters and receivers and it has a weight of $\approx 400g$, a size of $50 \times 50cm$ and a power consumption of 10 watts. Some drawbacks with an IR based solution is the requirement of a direct line of sight between the transmitting and the receiving modules and the cross interference between the sensors. Furthermore, these sensors are not suitable for MAVs operating in outdoor or large indoor environments due to their short working range.

1.3   Main contributions

This thesis proposes new solutions to address some of the challenges faced in the design and deployment of MAV teams for real missions, and in particular for search and rescue operations where teams of MAVs can be a crucial asset for fast recovery of victims and for saving human lives. We aim at proposing an on-board solution for swarm of MAVs, and novel methods for the three challenges of: target localization, relative positioning and self localization; where there is currently a lack of technological possibilities that could provide this information independent of any external systems while accommodating the strict constraints of micro aerial vehicles. We propose the novel concept of exploiting sound to obtain the required on-board sensor suite and strategies for addressing these challenges.

More specifically, this thesis contributes with:

- an on-board acoustic source localization system suitable for detection and localization of acoustic sources in the environment from airborne MAVs.

- proposing the innovative application of locating emergency acoustic sources from flying MAVs that could be crucial for search and rescue and/or aerial surveillance missions.

- novel methods for localization narrowband sound sources and overcoming the high ambiguity associated with localizing these type of sources.

- an on-board sensor suite and methods for allowing individuals in an MAV swarm to directly measure the relative bearing of other robots in the vicinity and estimators to estimate their relative locations, considering the current lack of technological possibilities that can be used on these small scale flying robots.

- a cooperative method for allowing individuals in a swarm of MAVs to obtain their three dimensional position, all relative to a single reference point, without the need of a communication network and independent of GPS and other external systems, while accommodating the motion constraints of flying robots.

The proposed solutions are all tested on actual flying robots in order to validate each method.

1.4 Objectives and thesis overview

The objectives of this thesis are closely related to the three challenges, described in the previous section, for designing truly autonomous MAV swarm, which also correlate with the thesis chapters.

- **Chapter 2: Localization of emergency sound sources on the ground**
  The first objective is to develop a strategy for providing individual MAVs with the ability to detect and localize narrowband sounds, such as emergency sound sources, on the ground with the aim of assisting rescuers in a search and rescue mission with fast localization and recovery of victims. This chapter initially describes an onboard sensor suite and methods for detection and measuring the direction of incoming sound waves. It then explains the limitations in localizing narrowband sounds and proposes multiple methods for tackling these limitations. It is shown how different sources of information, such as the Doppler shift in the sound frequency, caused due to the motion of the MAV, the motion dynamics of the flying robot, and the behaviour of the robot itself can be employed to obtain accurate localization estimates.

- **Chapter 3: Audio-based relative positioning for multiple micro air vehicle systems:**
  The second objective is to obtain an on-board solution for allowing individual robots to locate other swarm members in their vicinity. In this chapter, firstly, an on-board audio-based relative bearing measurement system is described. For this, a method that exploits the engine sound of other flying robots, in the absence of the self-engine noise, is proposed. Furthermore, the method is extended to achieve a longer detection range, to relax the self-engine noise constraint, and to identify the identity of the robots. A position estimator is then described to robustly estimate the relative range and bearings from the direction measurements. In this chapter, methods such as Fractional Fourier Transform(FrFT), time-delay of arrival direction estimation (TDOA), and particle filters are employed for sound source separation, relative direction measurement and relative position estimation, respectively.

- **Chapter 4: Audio-based localization for swarms of micro air vehicles:**
  The third objective is to propose a method for obtaining self-localization of individuals in an MAV swarm that is independent of any external systems. This chapter describes a method based on the cooperation of robots for obtaining the three-dimensional positioning information. Here it is shown how moving individuals can use their on-board audio based sensor suite to measure the relative direction of a sound-emitting beacon MAV and furthermore use these bearing-only measurements to obtain their three-dimensional positions throughout time and without the need of any communication network. In particular, we propose a solution that can accommodate the motion constraints of flying robots and we demonstrate it on fixed wing robots that are not able to stop or turn in-place and must always maintain a forward speed for remaining aloft.

- **Chapter 5: Conclusion:**
Finally, Chapter 5 summarizes and concludes the thesis. In this chapter, firstly the main accomplishments of this thesis work is described. Secondly, multiple potential applications for the described methods, and in general for swarms of micro aerial vehicles with hearing capabilities, is proposed. The chapter concludes by providing some possible future work towards improvement of the propositions provided in this work.

1.5   Publications during thesis work

Parts of this thesis was published in the following publications:

- Basiri, M., Schill, F., Lima, P., and Floreano, D. (2012). Robust acoustic source localization of emergency signals from micro air vehicles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4737 –4742

- Basiri, M., Schill, F., Floreano, D., and Lima, P. U. (2013). Audio-based relative positioning system for multiple micro air vehicle systems. In *Proceedings of Robotics: Science and Systems, Berlin, Germany*

- Basiri, M., Schill, F., Floreano, D., and Lima, P. U. (2014a). Audio-based localization for swarms of micro air vehicles. In *2014 IEEE International Conference on Robotics and Automation (ICRA), Hongkong*

papers submitted or under submission for possible publication:

- Basiri, M., Schill, F., Lima, P. U., and Floreano, D. (2014b). Localization of emergency acoustic sources by micro aerial vehicles. *Under review in IEEE Transactions on Robotics*

- Basiri, M., Schill, F., Lima, P. U., and Floreano, D. (2015). Onboard audio-based relative positioning system for teams of small aerial robots. *Under submission*

2 Localization of emergency sound sources on the ground

I<small>N</small> search and rescue operations, Micro Air Vehicles (MAV's) can assist rescuers to faster locate victims inside a large search area and to coordinate efforts of rescue teams. Acoustic signals play an important role in outdoor search and rescue missions. Personal alarms and safety whistles, as found on most aircraft life vests, are commonly carried by people engaging in outdoor activities, and are also used by rescue teams, as they allow transmission of a distress signal reliably over long distances and far beyond visibility. Such sources generate a loud narrowband sound signal, suitable for long range detections, that is difficult to localize by human listeners. For an MAV involved in a search and rescue mission, the ability to locate the source of a distress sound signal, such as the sound of an emergency whistle blown by a person in need of help, is significantly important and would allow fast localization of victims and rescuers during night time, through foliage and in adverse weather conditions such as dust, fog and smoke. In this chapter we present a real-time on-board sound source localization system for MAVs to autonomously locate emergency acoustic sources on the ground, such as the sound of a safety whistle or a personal alarm. We propose three different methods for localizing emergency sources with MAVs that are based on measuring the coherence between signals of four spatially separated on-board microphones. The first method involves designing an emergency sound source that allows immediate localization by the MAV while relying only on acoustic information. The other two methods allow localizing currently available emergency sources that are difficult to localize due to the high ambiguity associated with localization of narrow-band sounds. The second method uses a particle filter to combine information from the microphone array, the dynamics of the MAV, and the Doppler shift in the sound frequency caused due to the motion of the MAV. The third method involves actively controlling the robot's attitude and fusing acoustic measurements with attitude measurements for achieving accurate and robust estimates. Furthermore, we evaluate our methods in real world experiments where two types of flying micro air vehicles are used to locate and track a narrowband sound source on the ground.

2.1 Introduction

The main objective of a search and rescue mission is to quickly locate and extract victims from the disaster situation. A search effort in rough outdoor terrain can be very time consuming and physically challenging, and keeping track of the positions of multiple rescue teams in a large area without communication infrastructure can be an additional problem. Autonomous Micro Air Vehicles (MAVs) can assist rescuers to faster locate victims in a large search area, and help coordinate rescue efforts by reporting the location of rescue teams to the mission coordinator [Hauert et al., 2010]. They can directly reach potential target areas by flying over obstacles, cluttered and inaccessible terrains, and hence achieve area coverage faster than ground units. Furthermore, by employing a swarm of aerial robots and dividing potential target areas among them, it is possible to further increase the searching speed and obtain scans of large areas quickly. Locating human victims from aerial robots have raised the interest of few researchers over the recent years [Andriluka et al., 2010, Goodrich et al., 2008, Gaszczak et al., 2011, Reilly et al., 2010]. Almost in all related works, the detection of people has been investigated through images obtained from vision sensors.

Sound is one of the most important cues for locating people in a disaster situation. Sound waves travel in all directions and can be detected at long distances from the sound source, and beyond line of sight. During night time or in fog, dust, smoke, dense forests or otherwise cluttered environments, acoustic signals are far more reliable than visual cues. This is the main motivation behind the use of personal alarms and safety whistles in most survival kits and disaster preparedness supplies offered today. Safety whistles are an inexpensive and effective method for emitting a distress signal in emergency situations, and are also commonly used for basic signalling when noise or distance makes voice communication difficult. They are often used by people engaging in outdoor activities, such as hikers, mountaineers, skiers, boaters and scuba divers, and are commonly provided with airplane life vests. These whistles enable the user to generate a very loud and clear narrow-band sound (usually a pure tone between 2-5 kHz), which can be perceived distinctly from long ranges and in noisy environments, without making the signaller hoarse and exhausted. There is usually little environmental noise in the multi-kilohertz range, which enhances detectability. However, a disadvantage of high-frequency, narrow-band sounds is the difficulty for human listeners to correctly locate the direction of the source [Stern et al., 1988]. Similarly, personal alarms are capable of generating a loud high frequency narrowband sound signal with frequencies around the piezo element's resonance frequency. In addition, with the advancement of mobile phones and hand-held computers, applications for emitting distress signals through the device speakers is not out of reach, where a loud sound signal could be produced by focusing the power into a narrow frequency band.

The aim of this chapter is to develop an audio-based localization system for MAVs to locate and track emergency acoustic sources on the ground, while satisfying the strict constraints imposed on MAVs in terms of size, weight, processing power and electrical power consumption. Such a system could also provide a simple way of interaction between human operators and

MAVs. For example, a rescue team could easily signal its position to the mission coordinator by using specific whistle signals, or an operator could command robots to land in a desired spot. Other potential applications include aerial surveillance and environmental noise monitoring. In the next chapters we show how this system could further be exploited to address relative positioning and localization problems for the design of multi-MAV systems.

Hearing has always been one of the key senses among humans and animals allowing them to use sound for attracting attention, communication and localization purposes. Despite this, audition in robotics has not received great attention compared to vision, and most studies on this focus on speech recognition and localization of talkers for home, office, and humanoid robots [Matsusaka et al., 1999] [Okuno et al., 2002] [Nakadai et al., 2000]. In most works, a technique inspired by animal hearing called Inter-aural Time Difference (ITD) (also known as Time Difference of Arrival TDOA) is used for localizing sound sources. This method is based on measuring the time delay, caused by the finite speed of sound, between the signals perceived by two spatially separated microphones to estimate the direction of incoming sound waves. While the complex hearing capabilities of animals achieve good performance with only one pair of acoustic sensors, technical systems often use arrays of microphones for assisting robots in locating broadband sound sources in the environment [Valin et al., 2003]. In this chapter, we propose solutions relying on the TDOA method for localizing narrowband sounds and for designing an emergency source localization system for MAVs based on a small, compact and on-board array of low cost electret microphones. Design of new acoustic sensors that could also be used on MAVs have been investigated in some recent works [Ruffier et al., 2011] [de Bree et al., 2010].

In the TDOA sound source localization method, the time delay between signals of a microphone pair is generally estimated using cross-correlation. A problem faced with this approach is that it requires the sound source to be a broadband source [Buchner et al., 2005] and it fails to estimate the correct time delay for narrowband sounds. This is because with narrowband sounds there is an ambiguity in the time delay estimations which results in ambiguous direction estimation, i.e. coherence testing among the signals from different microphones no longer provides a unique time delay. This problem is particularly pronounced for higher frequencies.

We propose three different strategies for an MAV based emergency source localization system and for overcoming the problems with localizing such narrowband sound sources. A detailed explanation of each method is explained throughout this chapter.

- The first method consists of designing an emergency acoustic source, based on a piezo transducer, that generates a sound wave that can be localized immediately from the microphone signals. This is achieved by fast modulation of the sound frequency around the resonance frequency of the piezo element, for increasing the frequency bandwidth, and a modified TDOA estimation for obtaining a unique source direction.

15

The second and third methods investigate localization of existing emergency sources, or narrowband sources in general, where the TDOA method alone fails to correctly localize the source.

- The second method exploits the fast motion of the MAVs and the Doppler shift in the sound frequency, caused due to this motion, for resolving the ambiguous TDOA direction estimates. This is more suitable for fixed wing type of MAV's as they always need to maintain a forward speed for staying airborne.

- In the third proposed method we explain how in-place changing of the attitude by the flying robot could allow the robot to correctly locate narrowband sound sources. This is more suitable for rotorcraft MAVs such as quadrotors since they are capable of hovering and turning in place.

In both latter methods, our strategy is to use the ambiguous TDOA information along with other sources of information in order to obtain a more reliable estimate. For this, particle filters are employed, also known as sequential Monte Carlo method [Gilks et al., 1996]. Particle filtering is considered a powerful tool for handling localization, navigation and tracking problems [Doucet et al., 2001]. Few works show the effectiveness of particle filtering for tracking wide-band sound sources inside reverberant environments [Ward et al., 2003],[Asoh et al., 2004]. The work presented in this chapter focuses on the tracking of narrowband and single-frequency sound sources, proposes novel methods to incorporate acoustic informations and vehicle dynamics in a particle filter framework, and presents experimental results.

This chapter is organized as follows: Section 2.2 describes the TDOA localization method and explains the limitations in locating narrowband sound sources. The definitions and equations used throughout this chapter and parts of other chapters are described here. Section 2.3 explains the three proposed methods for an audio-based localization system for MAVs for the purpose of locating emergency acoustic sources. In Section 2.4 results of real experiments for each proposed methods are provided, where two types of flying MAVs are used to locate and track the location of emergency sound sources in the environment. Section 2.5 concludes the chapter, providing also clues for future work.

2.2   TDOA sound source localization

In TDOA localization an array of spatially separated microphones is used for estimating the 3D direction to a sound source. Due to finite speed of sound, incoming sound waves are picked up by the microphones at different times. By comparing the microphone signals and obtaining the time delay between them, it is possible to estimate the direction of arrival of the sound wave. TDOA localization can be divided into two main parts:

1. Coherence measuring
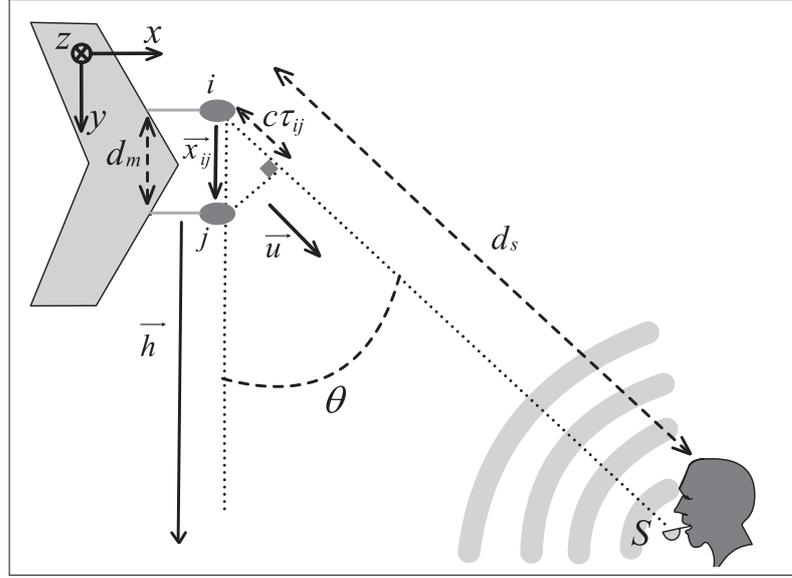
Figure 2.1: Illustration of a sound field consisting of two microphones $(i),(j)$ and a single sound source S. The microphones are placed on an MAV and d_m meters apart from one another. Vector u is a unit vector in the body-fixed coordinate system (x, y, z) pointing towards the sound source S.

2. Direction estimation

A detailed explanation of each part is provided throughout the rest of this section.

2.2.1 Coherence measuring

For a pair of spatially separated microphones i, j that experience sound waves from a common sound source S (as illustrated in Figure 2.1), a time delay τ_{ij} exists between their reception of the sound signal. This time delay is dependent on the angle θ between the sound wave-front and the microphone pair's baseline direction, and it can be approximated by considering the far field assumption ($d_s >> d_m$):

$$\tau_{ij} \approx \frac{d_m \cos\theta}{c} \tag{2.1}$$

where d_m is the distance between the microphones, c is the speed of sound. Thus the time delay τ_{ij} lies within the range:

$$-\frac{d_m}{c} < \tau_{ij} < \frac{d_m}{c} \tag{2.2}$$

Cross correlation is a commonly used technique for measuring the coherence and obtaining the time delay between two signals. Cross correlation between the digitized sequences of two

microphone signals $p_i[n], p_i[n]$, each having a length of N samples can be computed by

$$R_{ij}(\tau) = \sum_{n=0}^{N-1} p_i[n]\, p_j[n-\tau]$$

where $p_i[n]$ is the signal perceived by microphone i and τ is the correlation lag in samples in the range expressed by Equation (2.2). In order to reduce the computation time, the cross correlation function can be obtained in the frequency domain by computing the inverse Fourier transform of the cross spectrum:

$$R_{ij}(\tau) = \sum_{k=0}^{N-1} P_i[k]\, P_j^*[k]\, e^{i\frac{2\pi k\tau}{N}} \tag{2.3}$$

where $P_i(k)$ is the discrete Fourier transform of $p_i(n)$ and P_j^* denotes the complex conjugate of P_j. This results in a reduction of complexity from $O(N^2)$ to $O(N\log N)$, making it more suitable for on-line computations.

The correlation value $R_{ij}(\tau)$ reaches a maximum value at $\tau = \tau_{ij}$. The shape of the cross correlation R_{ij} is dependent on the statistical properties of the sound signal itself. If the sound source is an ideal white noise then R_{ij} is equal to an impulse function transposed by τ_{ij}. However, if the sound source is a band limited white noise, R_{ij} no longer has an impulse shaped peak but instead it has a more broadened peak centred at τ_{ij}. The actual shape of the correlation R_{ij} can be described mathematically by [Ferguson, 1999]

$$R_{ij}(\tau) = P_a \frac{\sin(\pi B(\tau - \tau_{ij}))}{\pi B(\tau - \tau_{ij})} \cos(2\pi f_0(\tau - \tau_{ij})) \tag{2.4}$$

where B is the bandwidth of the signal, f_0 is the centre frequency of the bandwidth, and P_a is the signal power. As the bandwidth approaches zero, i.e. the sound source is a pure sinusoid with frequency f_0, the correlation R_{ij} becomes a periodic sine wave with peaks of the same amplitude and period of $1/f_0$. This makes the identification of the peak corresponding to the correct time delay ambiguous. Figure 2.2 shows cross-correlations in a simple sound field case explained in Figure 2.1, for three different sound sources. It can be seen that the cross correlation of narrowband sounds from a whistle and a piezo alarm contain multiple peaks; hence making it impossible to identify the true time delay.

One method to tackle the time delay ambiguity problem for narrow-band sound sources is to simply decrease the distance between the microphones in order to avoid multiple cross correlation peaks. To obtain a single cross-correlation peak, the microphone pair's inter-distance must satisfy

$$d_m < \frac{c}{2f_0} \tag{2.5}$$

However, the angular resolution of a microphone array for a given sampling frequency decreases for smaller distances between the microphones and as we are interested in locating
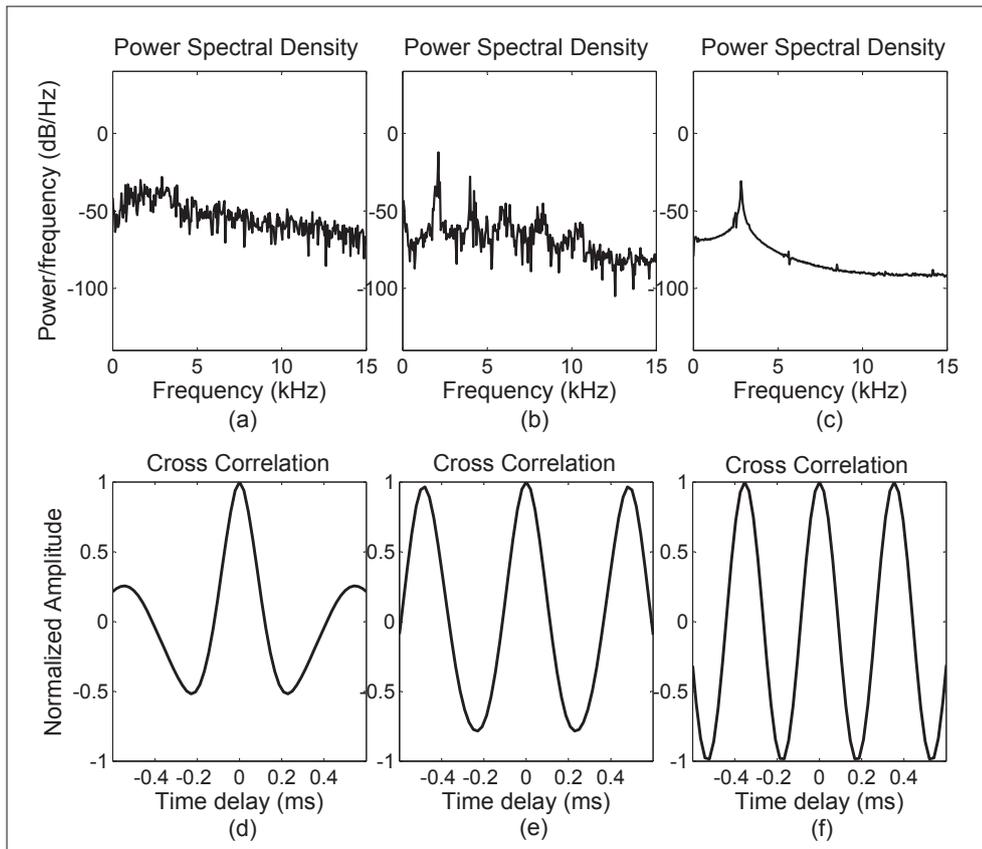
Figure 2.2: Cross-correlations for three different sound sources. The power spectral density of every sound source is shown in graphs(a-c) and their corresponding cross correlation is shown below each graph (d-f). The three sound sources are: (a) air blower (b) emergency whistle (c) piezo alarm

high frequency sounds, this would lead to poor estimations.

2.2.2 Direction Estimation

We explained previously how cross correlation could be used to obtain a measure of similarity between signals of a microphone pair. Here, we will explain how the measure of similarity for an array of microphones could be used along with knowledge of the microphone array's geometry to estimate probable sound source directions.

Let's consider the sound field scenario in Figure 2.1 and let's initially assume that the sound source S is a broadband source that leads to a unique time delay τ_{ij} when correlating signals of microphones $(i),(j)$. Considering the far field assumption ($d_s >> d_m$), and using Equation (2.1), it is possible to approximate:

$$\cos\theta = \frac{c\tau_{ij}}{d_m} \tag{2.6}$$

Also, from the cosine law it is possible to derive:

$$\cos\theta = \frac{\overrightarrow{x}_{ij}.\overrightarrow{u}}{d_m} \tag{2.7}$$

where $\overrightarrow{u}$ is a unit vector pointing towards the sound source. By combining Equations (2.6) and (2.7) we get

$$\overrightarrow{x}_{ij}.\overrightarrow{u} = c\tau_{ij} \tag{2.8}$$

In three dimensional space and by considering a body-fixed coordinate system that is attached to the MAV's frame, the vectors $\overrightarrow{u}$ and $\overrightarrow{x}_{ij}$ are defined as

$$\begin{aligned}
\overrightarrow{u} &= (u_x, u_y, u_z) \\
\overrightarrow{x}_{ij} &= (m_{ix} - m_{jx}, m_{iy} - m_{jy}, m_{iz} - m_{jz}) \\
&= (m_{ijx}, m_{ijy}, m_{ijz})
\end{aligned} \tag{2.9}$$

where (m_{ix}, m_{iy}, m_{iz}) are the coordinates of microphone i. To solve for a unique 3D direction $\overrightarrow{u}$, from (2.8), a minimum of three microphone pairs that are not all located on the same plane is required.

A simple and powerful method for estimating the source direction from the coherence measurements, known in the literature as Global Coherence Field(GCF)[DiBiase et al., 2001] is to simply compute the sum of cross correlation values, of all the microphone pairs, for every plausible direction $\overrightarrow{u}$.

$$B(\overrightarrow{u}) = \sum_{\forall pairs} R_{ij}\left(\tau_{ij}(\overrightarrow{u})\right) \tag{2.10}$$

where $\tau_{ij}(\overrightarrow{u})$ is the expected time delay for pair i, j, if the source was in direction $\overrightarrow{u}$, and is computed from Equation (2.8). In case of broadband sources, the direction of the source can be derived as the direction that maximizes $B(\overrightarrow{u})$.

An alternative method for computing the direction towards a broadband source is to firstly find the τ_{Mij} that maximizes R_{ij} and then to use the least square search method to compute the direction that best fits the set of time delays for all pair combinations. From (2.8) and (2.9) a linear system of N equations can be obtained when N different microphone-pairs are used:

$$Au = b \tag{2.11}$$

$$A = \begin{bmatrix} m_{12x} & m_{12y} & m_{12z} \\ m_{13x} & m_{13y} & m_{13z} \\ m_{23x} & m_{23y} & m_{23z} \\ \vdots & \vdots & \vdots \\ m_{ijx} & m_{ijy} & m_{ijz} \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ N \end{matrix} \quad u = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} \quad b = \begin{bmatrix} c\tau_{M12} \\ c\tau_{M13} \\ c\tau_{M23} \\ \vdots \\ c\tau_{Mij} \end{bmatrix} \tag{2.12}$$

The system of Equations (2.11) can be used to compute a unique solution u if three or more microphone-pairs exist which are not all on the same plane. A value for u that best satisfies the system of Equations (2.11) can be estimated using the linear least square method [Lawson and Hanson, 1995]

$$u_{ls} = (A^T A)^{-1} A^T b \tag{2.13}$$

to minimize the sum of the squared errors

$$\varepsilon_r = \sum (A u_{ls} - b)^2 \tag{2.14}$$

Both explained methods for direction estimation provide a unique solution in the case of a broadband sound source as there is a single global maximum in the correlation for each pair. The least square (LS) method is less computationally complex and hence more suitable for real-time applications. However, as it only considers the maximum correlation point for every microphone pair it is considered to be the weaker approach and its accuracy is shown to drop for low signal to noise situations [Brutti et al., 2008a]. Throughout this work both methods were used. The LS method was chosen initially for fixed-wing type robots as their fast speed requires a faster solution and also since by reducing their engine power and gliding, they are capable of increasing the signal to noise ratio. The GCF was used in rotor craft type of robots as they suffer more from engine noise.

In Section 2.2.1 it was shown that coherence measuring for narrowband sounds leads to multiple peaks in the cross correlations. This also results in multiple global maximum in the sum of cross correlation value $B(\vec{u})$, in Equation (2.10), and hence multiple potential directions to the sound source could be interpreted. Figure 2.3 shows the normalized values of $B(\vec{u})$ for a narrowband sound of 5kHz obtained with a tetrahedral microphone array of length 18 cm and for all feasible directions $\vec{u}$. Furthermore, the LS method cannot be used as there are more than one time delay that maximizes each cross correlation. In later sections we describe methods for overcoming this problem. Our strategy is to take into account all potential directions and fuse this information with other sources of information available to obtain the correct direction. In Section 2.3.2.2 the value of $B(\vec{u})$ is directly used as the likelihood that direction $\vec{u}$ is pointing towards the sound source. In Section 2.3.2.1, the LS method is used instead to derive a set of potential source directions. This is achieved by
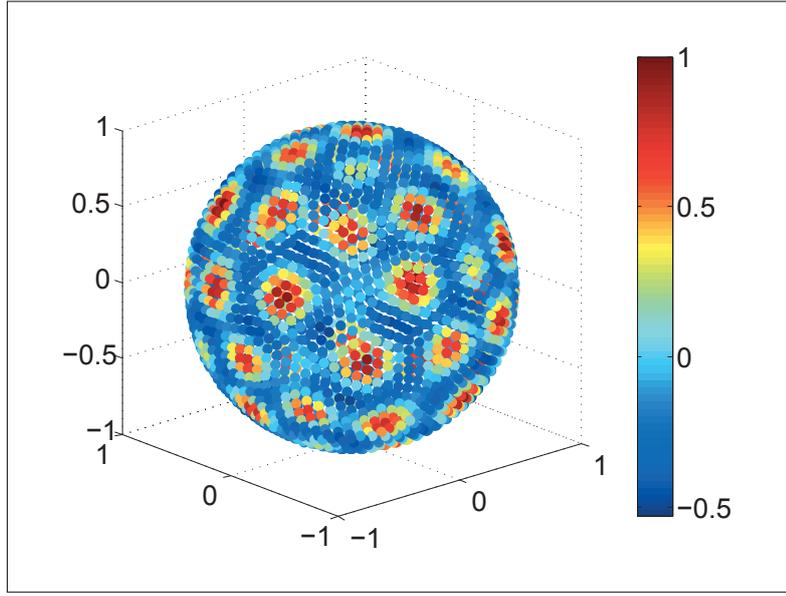
Figure 2.3: Normalized values of $B(\vec{u})$, shown by colors, for a narrowband sound of 5kHz obtained with a tetrahedral microphone array of length 18 cm. Each cell on the unit sphere represents a direction that is defined by a vector starting from the origin and ending on the cell.

computing the source direction u_{ls} and error ε_r, for all possible time delay combinations using Equations (2.13) and (2.14). Those values of u_{ls} that have the lowest errors are considered as potential directions towards the sound source. Hence, a set of possible source directions is obtained:

$$\{u_k, k = 1 : M\} \tag{2.15}$$

where M is equal to the number of corresponding peaks among microphone pairs.

2.3 Localization of emergency acoustic sources

In the previous section we described how an on-board microphone array could be used to obtain information about the direction of a sound source relative to an MAV. We furthermore explained the limitations of obtaining a unique and accurate direction estimate for existing emergency acoustic sources emitting a narrowband sound signal. In this section we will explain the three proposed methods for having an audio-based localization system, suitable for MAVs, for the purpose of locating emergency acoustic sources.

2.3.1   Design of an emergency acoustic source

A simple and effective solution for designing an emergency source localization system for MAVs is to simply design an acoustic source that facilitates the TDOA localization process by resulting in a unique time delay when measuring the coherence between the microphone signals. This would allow the robot to obtain a unique and instantaneous bearing measurement to the target while only relying on acoustic information. Multiple criteria were considered for designing such a source. Similar to other emergency acoustic devices, it should be easy to use and to carry by human operators. It has to generate a loud sound that can be picked up from long distances by both human ears and the MAV's on-board microphones. The sound's frequency should also be in the multi-kilohertz range to enhance detection against low frequency environmental, wind and the engine noise of the robots. For this, a piezo based device was developed. Piezo transducers are simple, inexpensive and lightweight devices that are already used in most personal alarms. Piezo transducers generate sound by converting electrical pulses into mechanical vibrations. The resulting sound can be very loud if the frequency of the vibrations are close to the resonance frequency of the piezo element. To generate a loud sound wave that is required here and furthermore to obtain a unique global maximum in the cross correlations, a driving circuit based on a micro controller was designed to produce a continuous and band-limited periodic linear chirp signal around the piezo's resonance frequency. The instantaneous frequency of this signal is computed from

$$f(t) = F_H - \frac{F_H - F_L}{\Delta T} \, \mathrm{mod}\left(\frac{t}{\Delta T}\right) \tag{2.16}$$

where F_H and F_L are the highest and lowest frequencies of the chirp, that is found empirically as the limits above and below the piezo's resonance frequency where the output sound power is still above a certain threshold which is suitable for long range detections. ΔT is the duration of a single chirp that is chosen to be equal to the length of the time window used in the coherence measuring, i.e. $\Delta T = N/F_s$ where F_s is the sampling frequency and N is the number of samples used for coherence measuring. Figure (2.4) shows the spectrogram of a sound recording from this source illustrating the change in the sound frequency.

On the perceiving end, a template of the chirp that is stored in the memory is used by the robot to detect the presence of this sound source in the environment. This is achieved by continuously cross correlating the template signal with one of the microphone signals. When a good coherence is detected, it can be concluded that the sound is from the desired source. Once the sound is detected, the TDOA method explained previously is employed to estimate the direction towards the sound source. However, as the output sound power is not uniform for all the frequencies of the chirp and the sound is also significantly louder at the resonance frequency, it results in a non equal contribution of the frequencies when computing the cross correlation using Equation (2.3). This could lead to multiple wide correlation peaks. A weighting function was introduced into Equation (2.3) by [Knapp and Carter, 1976] in order to
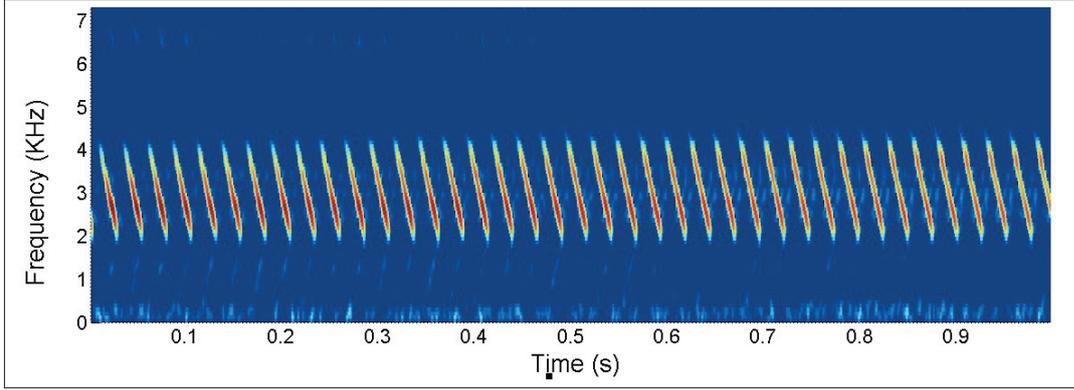
Figure 2.4: Spectrogram of one second of sound recoding from the proposed piezo based emergency source. The individual chirp duration ΔT is equal to $\Delta T = N/F_s$, where N is the number of samples used in the coherence measuring and F_s is the sampling frequency. In this work $N = 1024$, $F_s = 40$kHz and $\Delta T = 25.6ms$

solve the problem of wide cross correlation peaks in broadband sound localization. This is achieved by whitening the cross-spectrum of the signals and allowing equal contribution of all frequencies in the cross correlation. A modified version of this weighting function is used here instead to allow dampening the resonance frequency and providing equal contribution of the entire chirp frequencies:

$$R_{ij}(\tau) = \sum_{k=0}^{N-1} \chi \left[\frac{P_i P_j^*}{|P_i||P_j|} \right] e^{i \frac{2\pi k \tau}{N}} \tag{2.17}$$

$$\chi = \begin{cases} 1 & f_L < f < f_H \\ 0 & \text{otherwise} \end{cases}$$

Result of experiments using this method is presented in Section 2.4.1.

2.3.2 Localizing ambiguous narrowband sources

The previous section described the developing of a new sound source and the TDOA based localization system that results in immediate localization by the MAVs depending entirely on sound waves. However, it is of interest that MAVs can also detect and locate the currently available narrowband emergency sources, such as whistles or personal alarms, that lead to ambiguous TDOA measurements. In this section we describe two different estimators for locating such sources. Both estimators employ particle filters for fusing TDOA measurements with other sources of information in order to recursively estimate the probability density of the target location. Information sources such as the behaviour of the MAV in time, provided by on-board proprioceptive sensors, TDOA measurements available up to the current time and the Doppler shift in the sound frequency is used to obtain reliable and accurate direction estimates. The general algorithm of the proposed particle-based estimators [Doucet et al.,

1. Initiate a set of N particles with equal weights:
$S_i(0) = \{(\vec{u_i}(0), w_i(0) = 1/N_p) : i = 1, 2, .. N_p\}$

2. Repeat for every data frame:

   (a) **Prediction:** Predict a new set of particles $\tilde{S}_i(t)$ from $S_i(t-1)$ using the probabilistic motion model of the MAV and on-board proprioceptive sensors.

   (b) **Update:** Update the weight w_i of every particle by finding the likelihood of acquiring audio-based measurements $y(t)$ given the particle's predicted state. $w_i(t) = p(y(t)|\tilde{u}_i(t))$.

   (c) Normalize w_i to have $\sum\limits_{i=1}^{N} w_i = 1$.

   (d) Estimate the target direction, and a reliability measure for this estimation, from the probability density function represented by the particle set.

   (e) Form particle set $S_i(t)$ by re-sampling $\tilde{S}_i(t)$ according to the weights of particles.

Table 2.1: General algorithm of the proposed particle-based estimator for estimating the direction of ambiguous narrowband sounds

2001], illustrated in Table 2.1, is described here. More details of each estimator is provided in the following subsections.

At time instant t, all the hypotheses about the target's direction is modelled using a set of N_p particles of direction vectors $\vec{u_i}(t)$ and weight $w_i(t)$:

$$S_i(t) = \{(\vec{u_i}(t), w_i(t)) : i = 1, 2, .. N_p\} \tag{2.18}$$

where $\vec{u_i} = (u_{xi}, u_{yi}, u_{zi})$ is a unit vector in the body-fixed coordinate system that starts at the origin and points towards a direction. $\vec{u_i}$ can also be described in the body-fixed spherical coordinate system $(r, \angle\phi, \angle\theta)$ by:

$$\vec{u_i} = (1, \phi_i, \theta_i) \qquad i = 1, 2, .. N_p \tag{2.19}$$

where ϕ_i is the azimuth defined in the range $[-\pi, \pi]$ and θ_i is the elevation defined in the range $[-\pi/2, \pi/2]$. The estimators start by forming an initial set of N_p particles $S_i(0)$ with uniform weights $w_i(0) = 1/N_p$. Particles either could be generated uniformly over the entire state space, or only over a desired part of the state space if some prior knowledge about the possible location of the target is available. In the proposed problem of localizing targets on the ground from a flying MAV the initial state space is reduced to all vectors pointing towards the ground.

Prediction and Update are the two main recursive steps of the particle based estimators. In the

prediction step, a set of new particles $\tilde{S}_i(t)$ is predicted by propagating the states of $S_i(t-1)$ according to a probabilistic motion model. This is achieved by transforming the vectors $\vec{u}_i(t-1)$ to $\vec{u}_i(t)$ using the information gathered from the MAV's on-board proprioceptive sensors indicating the change in the state of the MAV. In the update step, acoustic measurements are employed to investigate the likelihood of obtaining these measurements for all particles, and particles are then weighted according to this measure. Furthermore, at each time step, particles are re-sampled according to their weights and a direction to the target is estimated from the probability density function represented by the particle set. Detailed description of the mentioned steps for both proposed estimators are provided in the following subsections.

2.3.2.1 Exploiting the motion of the MAV

The first estimator was initially designed to address the problem of emergency source localization using fixed wing type of MAVs. These MAVs always need to maintain a forward motion in order to stay airborne. The idea here is to make use of this motion to obtain a unique and accurate direction estimate for narrowband sounds. Due to the relative speed between the MAV and the target, the perceived sound's frequency is different from the source's frequency. This is known as the Doppler effect. The relationship between the frequency f_0 of the sound source and the observed frequency f is defined by the equation:

$$f = f_0 \frac{c + v_o}{c + v_s} \tag{2.20}$$

where c is the speed of sound, v_o and v_s are the components of the observer and the sound source velocities that is relative to each other. With the assumption of knowledge about the source frequency f_0, Equation (2.20) can be used to obtain a measure of the relative speed $v_r = (v_o - v_s)$ between a moving MAV and the stationary sound source $v_s = 0$:

$$v_r = c(\frac{f}{f_0} - 1) \tag{2.21}$$

where f is measured by searching for the maximum peak in the power spectral density of the sound measurements inside the range $f_0 - f_m < f < f_0 + f_m$, with f_m being the maximum possible shift in frequency. f_m is obtained from the absolute speed v_R of the MAV

$$f_m = \frac{v_R}{c} f_0 \tag{2.22}$$

Furthermore, existence of a peak within the defined range is used for detecting the presence of a source in the environment prior to executing the localization algorithm.

The model used for the **prediction step** of the estimator assumes that the robot has only forward motion (i.e. along the x axis on the body-fixed coordinate system), and the target is located on the ground, i.e. at zero height(see Figure 2.5). The prediction procedure starts by estimating a target distance d_i for every particle based on the MAV's current height and
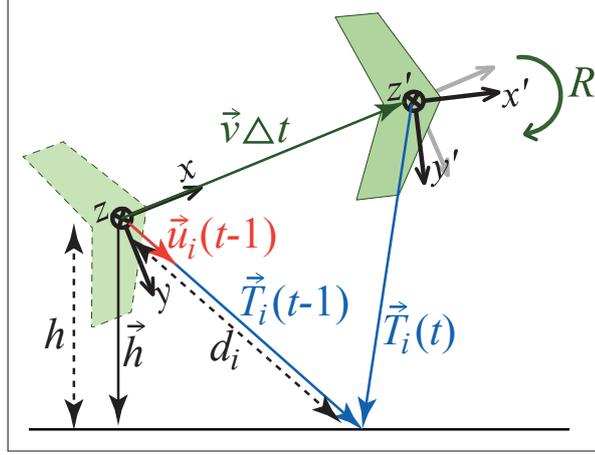
Figure 2.5: Position of a robot in two successive time steps along with the vectors used in the prediction step of the particle based estimator.

orientation:

$$d_i = \frac{h}{\cos \angle (\overrightarrow{u}_i(t-1), \overrightarrow{h})} \tag{2.23}$$

where $\angle(\overrightarrow{u}_i(t-1), \overrightarrow{h})$ is the shortest angle between vectors $\overrightarrow{u}_i$ and the height vector $\overrightarrow{h}$. From (2.19) and (2.23), it is possible to construct a vector T_i that extends the unit vector $\overrightarrow{u}_i$ until it reaches the ground.

$$\overrightarrow{T}_i(t-1) = (d_i, \phi_i, \theta_i) \tag{2.24}$$

This vector is then propagated by taking into account the translation and rotations of the body-fixed coordinate system due to the speed $\overrightarrow{v}$ and the change in the yaw ($\Delta\lambda$), pitch ($\Delta\beta$) and roll ($\Delta\alpha$) angles of the MAV.

$$\overrightarrow{T}_i(t) = R(\Delta\lambda, \Delta\beta, \Delta\alpha)(\overrightarrow{T}_i(t-1) - \overrightarrow{v}\Delta t) \tag{2.25}$$

where Δt is the time interval between the updates and R is a rotation matrix representing the rotations of the MAV. Finally, to account for the uncertainty in the predictions, a vector $\overrightarrow{u}_i$ is randomly generated from the set of all possible direction vectors with constraint:

$$\angle(\overrightarrow{T}_i, \overrightarrow{u}_i) \sim N(0, \sigma_A) \tag{2.26}$$

where $N(0, \sigma_A)$ is a normal distribution with mean zero and standard deviation σ_A. The value of σ_A is chosen in relation with the accuracy of the model that is used. The random 3D vector $\overrightarrow{u}_i$, that satisfies constraint (2.26), is obtained by simply generating a random vector $\vec{e}$, relative to the direction of the z axis and with distribution $\angle(\overrightarrow{z}, \overrightarrow{e}) \sim N(0, \sigma_A)$, and then rotating it so

27

that it is relative to the vector $\overrightarrow{T}_i$:

$$\tilde{\overrightarrow{u}}_i = \left(R_y(\frac{\pi}{2})R_x(-\phi_T)R_y(-\theta_T) \right)\vec{e} \tag{2.27}$$

where (ϕ_T, θ_T) are the azimuth and elevation of vector $\overrightarrow{T}_i(t)$, $R_a(b)$ is the basic rotation matrix for rotating vectors by angle "b" around the axis "a", and $\vec{e}$ is defined in spherical coordinate system as:

$$\vec{e}_{(r,\phi,\theta)} = \left(1, e_1, \frac{\pi}{2} - e_2 \right) \tag{2.28}$$

where $e_2 \sim N(0, \sigma_A)$ and $e_1 \sim U(-\pi, \pi)$ is a uniformly generated number in the range $(-\pi, \pi)$.

For the **update step** of the estimator we employ the set of M direction measurements from the TDOA localization unit (2.15) and the relative speed measurement v_r from the Doppler speed estimation (2.21) to investigate the likelihood of obtaining these measurements for every particle. Particles are then weighted according to this measure. For this investigation, we propose the likelihood function:

$$w_i = e^{\left(-\varepsilon_{Di}^2 / 2\sigma_D^2 \right)} \sum_{k=1}^{M} e^{\left(-\varepsilon_{Tik}^2 / 2\sigma_{Ti}^2 \right)} \tag{2.29}$$

where ε_{Tik} is the shortest angle between vector $\overrightarrow{u}_i$ and the kth TDOA measurement $\overrightarrow{u}_k$, and ε_{Di} is the error between the computed relative speed $\tilde{v}_{ri}$ and the measured relative speed value v_r:

$$\begin{aligned} \varepsilon_{Tik} &= \angle(\overrightarrow{u}_k, \overrightarrow{u}_i) \\ \varepsilon_{Di} &= |v_r - \tilde{v}_{ri}| \end{aligned} \tag{2.30}$$

where the computed relative speed $\tilde{v}_{ri}$ is simply calculated using the MAV's velocity vector $\overrightarrow{V}_R$ and the vector $\overrightarrow{u}_i$:

$$\tilde{v}_{ri} = \left\| \overrightarrow{V}_R \right\| \cos\angle(\overrightarrow{V}_R, \overrightarrow{u}_i) \tag{2.31}$$

The values of σ_T and σ_D in (2.29) reflect the confidence of Doppler-speed and TDOA measurements respectively and are found empirically.

The direction to the target is estimated at each time step from the probability density function represented by the particle set. For this a weighted mean of all particles' states could be used. However, to avoid inaccurate estimations for situations with multi-modal distributions, a weighted mean of particles located in a local neighbourhood of the particle with the highest weight is used instead:

$$\bar{u}_T = \sum_{i=1}^{K} w_i \overrightarrow{u}_i \qquad \forall \{\overrightarrow{u_i}\} \in S_i : \angle(\overrightarrow{u}_i, \overrightarrow{u}_{max}) < \xi \tag{2.32}$$

The experiments and results corresponding to this section is described in Section 2.4.2.

2.3.2.2 Behaviour based localization

Another method of narrowband sound localization is proposed in order to allow in-place localization that is suitable for rotorcraft type of MAVs such as helicopters and quadrotors. Such a solution would be advantageous for these type of vehicles as they are able to maintain their position once the sound is detected to allow a constant signal-to-noise ratio and avoid the loss of this detection. Furthermore, these type of MAVs are capable of flying in both indoors and outdoors which is advantageous in many rescue operations. Inspired by the ability of some animals in improving sound localization with the aid of head movements [Populin, 2006], we hypothesized that controlling the behaviour of the robot could allow elimination of ambiguities related with narrowband sound localization. The idea with this method is that once the robot detects the presence of a sound source, by in-place platform rotations it can obtain the true target direction. An advantage of this method over the former method is that no prior knowledge about the sound source frequency is required.

Once again let's consider the sound field scenario of Figure 2.1 where a single microphone pair is in the presence of a high frequency narrowband sound source of frequency f_0. The actual time delay τ_{ij} between the two signals is a function of the angle θ with the relation expressed in Equation (2.6). However, as previously explained, for a narrowband sound, coherence measuring results in multiple potential time delays, τ_p, inside the range expressed by Equation (2.2):

$$\tau_p = \tau_{ij} + \frac{n}{f_0} \qquad \forall n \ \in \mathbb{Z} : \left| \tau_p \right| \leq \frac{d_m}{c} \tag{2.33}$$

From Equations (2.33) and (2.6), the set of potential angles θ_p when the sound source is at angle θ can be obtained:

$$\theta_p = \cos^{-1}\left(\cos\theta + \frac{cn}{d_m f_0} \right) \qquad \forall n \ \in \mathbb{Z} : \left| \cos\theta_p \right| \leq 1 \tag{2.34}$$

Figure 2.6 illustrates a plot of the potential angles θ_p, computed using Equation (2.34), for different source angles θ when a sound source of 5kHz and microphone distance of 0.21 meters is used. An ambiguity in the form of 6-7 possible directions for every angle exists. However, it can also be noted that a linear change in the angle θ results in a linear change in only the potential direction corresponding to the true source direction. This is used as the main concept behind the proposed behaviour based localization method. Once the robot detects the presence of an ambiguous sound source in the environment, by changing the orientation of its on-board microphone array and comparing the acoustic measurements with the attitude information that is provided by its on-board sensors, it can eliminate the ambiguous directions and obtain the true source direction.
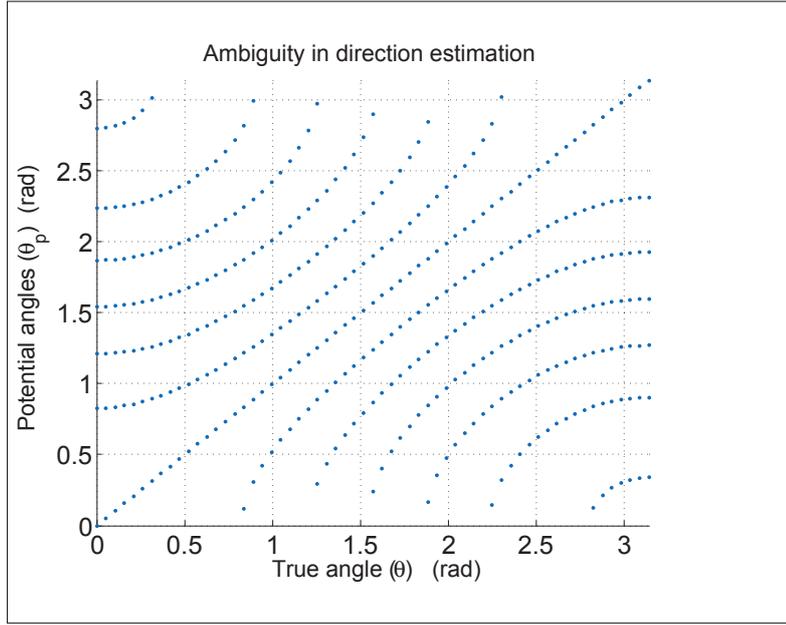
Figure 2.6: Plot of potential TDOA angles towards a narrowband sound source of 5kHz for different source angles θ relative to a microphone pair with inter distance of 0.21 meters. Multiple potential angles exist for every direction due to the ambiguity in the time delay estimations.

A particle based estimator is derived to fuse the attitude information with the acoustic TDOA measurements and to estimate the target's direction throughout time. The state space defined in Equation (2.19) is specified for every particle so that each particle represents a 3D direction in the body fixed coordinate system. Initially N_I particles are distributed uniformly over the part of the state space where it is pointing towards the ground. Figure 2.7 shows two representation of the initial particle distribution in a body fixed Cartesian and in a 2D plot of azimuth versus elevation. To have a uniform distribution over the entire potential state space, particles are initially placed on the corners of a geodesic grid on the surface of a unit sphere in the Cartesian coordinate system.

Upon first detection of the sound, the particles are weighted based on the sum of cross correlations $B(\vec{u}_i)$ from the TDOA measurements. For this, the time delay $\tau(\vec{u}_i)$ is initially computed from Equation (2.8). Furthermore, Equation (2.10) is used to compute the sum of cross correlations $B(\vec{u}_i)$ where cubic-spline interpolation is used to interpolate the values of $R_{ij}(\tau_{ij}(\vec{u}))$ from the discrete cross correlations R_{ij}. A new set of N_p particles, with $N_p < N_I$, is drawn from this initial distribution proportional to their weights. The robot then initiates an in-place rotation around its local z-axis, i.e. in yaw direction, and iteratively performs the prediction and update steps of the estimator.

In the prediction step, firstly particles are propagated according to the change in the robot's attitude compared to the previous state. The change in the yaw ($\Delta\lambda$), pitch ($\Delta\beta$) and roll ($\Delta\alpha$)
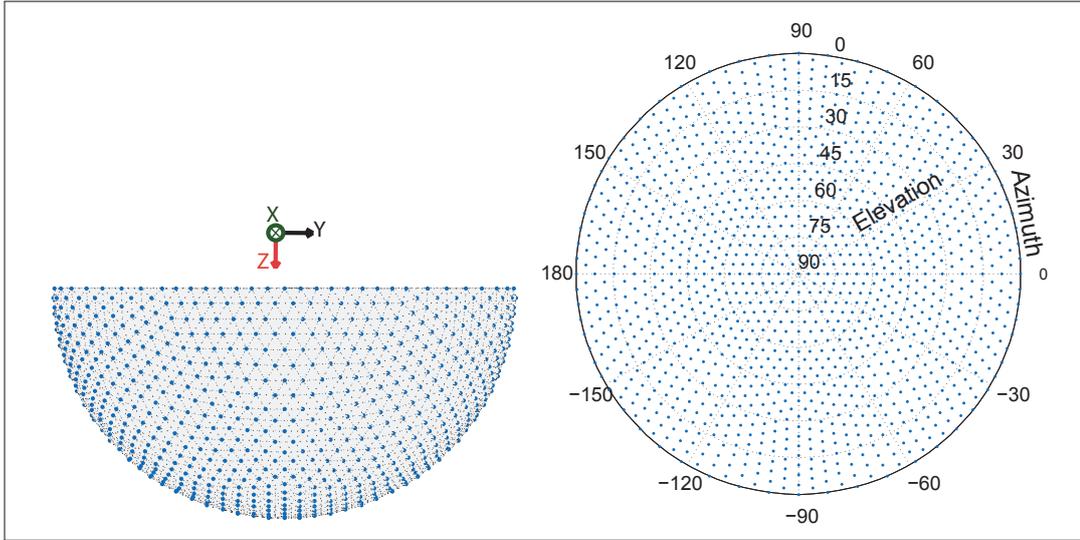
Figure 2.7: Plots of initial particle distribution relative to the body fixed coordinate system where each particle defines a 3D direction. Left) Plot of particles in Cartesian coordinate system. Particles are uniformly placed on the surface of a half sphere that represents the directions towards the ground. Right) Plot of azimuth versus elevation

attitude angles, obtained from on-board sensors, are used to propagate the particles:

$$\overrightarrow{u_i}(t) = R(\Delta\lambda, \Delta\beta, \Delta\alpha)(\overrightarrow{u}_i(t-1)) \tag{2.35}$$

Furthermore, similar to Equation (2.26), a set of predicted particles $\tilde{u}_i$ is generated by changing each particle $\overrightarrow{u}_i$ by a random angle that is drawn from a normal distribution with a standard deviation relative to the uncertainty in the attitude measurements. This random change in the directions also allows the exploration of the particles beyond their initial grid formation which leads to a finer search of the optima in the sum of cross correlation function B. In the update step, predicted particles $\tilde{S}_i(t)$ are once again weighted based on the sum of cross correlation values $B(\overrightarrow{u}_i)$. Finally, particles are re-sampled according to their weights. A systematic re-sampling approach [Arulampalam et al., 2002] that is suitable for real time implementation is used in this work.

Figure 2.8 illustrates the change in the particle distribution in an experiment where a constant in-place heading rotation of the robot results in the convergence of the particles to the correct source direction. The uncertainty existing in the TDOA measurements can be seen in the initial particle distribution after the first update. This uncertainty is reduced throughout time and a unique direction is estimated after a gradual decrease in the heading.

After each iteration, an estimated target direction $\vec{u}_T$ and a reliability factor X are estimated from the particle distribution. These are obtained using a simple vector addition of all the
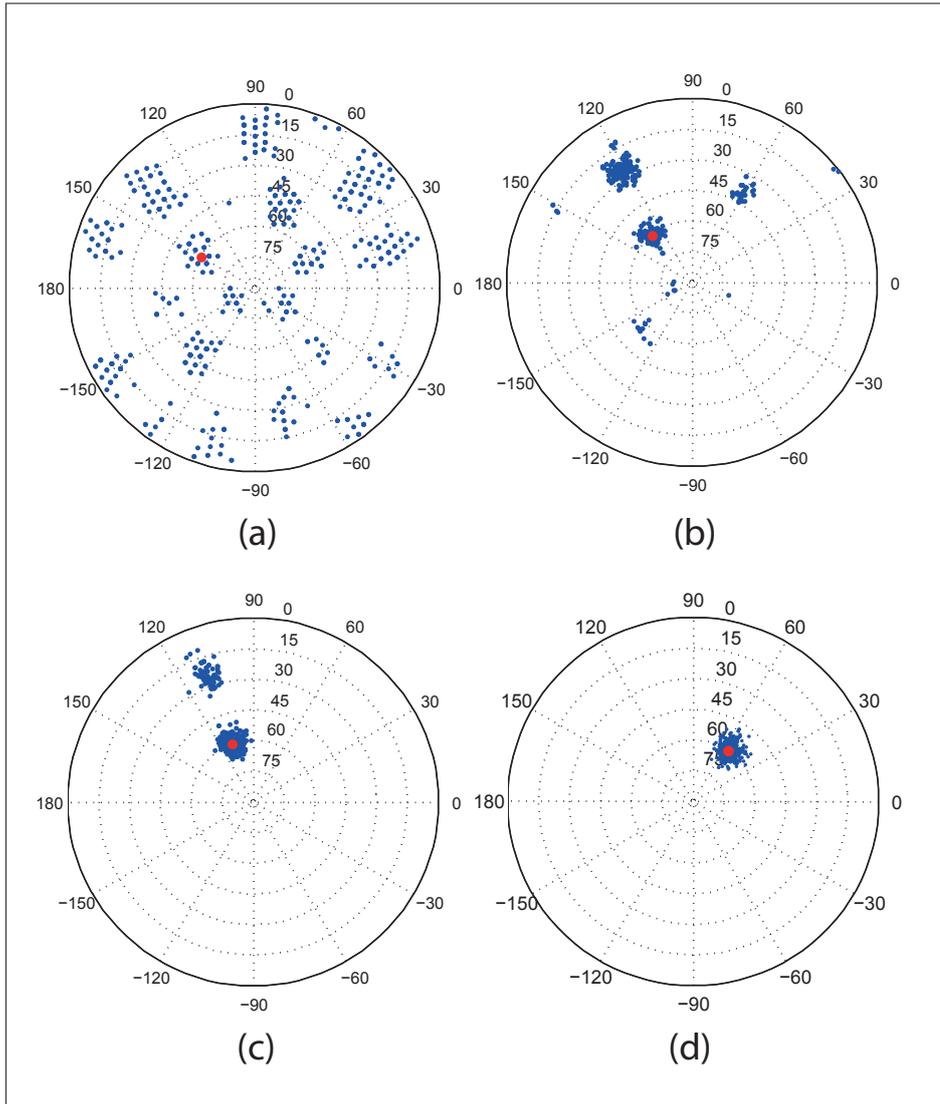
Figure 2.8: The change of the particle distribution in an experiment where the rotorcraft MAV of Figure 2.9 is rotated in yaw to correctly locate a sound source of 5kHz. a) After initial detection of sound b) After $\approx 20^o$ rotation c) After $\approx 40^o$ d) After $\approx 95^o$

particles, hence resulting in a fast computation that is suitable for real time implementations.

$$\vec{U}_T = X\vec{u}_T = \sum_{i=1}^{N_P} \vec{u}_i \tag{2.36}$$

where $\vec{u}_T$ is a unit vector along the direction of the total vector $\vec{U}_T$ and X is the norm of vector $\vec{U}_T$. The value of X reflects the disparity of the particles and is used as a measure of reliability for the estimated direction. The closer the value of X is to the particle number N_p, the more concentrated is the particle distribution. In this work the value of X is used by the robot to detect the convergence of the particles to a unique solution. The experiments and results

related to this section is described in Section 2.4.3.

2.4 Experiments and results

To test and verify the proposed methods, multiple experiments were performed with the two MAV platforms shown in Figure 2.9. A microphone array consisting of four microphones was mounted on both MAVs. The minimum number of four microphones, required for 3D direction estimation, was used to minimize the hardware and computational loads. The positions of the microphones were mainly chosen according to the constraints of the MAVs while trying to have an array shape close to a tetrahedron in order to have approximately equal performance in every direction [Hu et al., 2011]. For the fixed wing MAV, the microphones were positioned to form a regular tetrahedron of edge length 10 cm. For the quadrotor MAV, three microphones forming a triangle of edge length 18cm were placed between the propellers to stop the propeller airflow from influencing the microphones and the fourth microphone was placed under the MAV. The MAVs were equipped with autopilots that allowed them to fly fully autonomously to predefined waypoints.

The measurements required for the prediction steps of the estimators were obtained entirely using on-board sensors. For the fixed wing MAV, the change in roll and pitch of the MAV were measured using on-board gyroscopes and the airspeed and altitude using an absolute and a differential pressure sensor. As no compass was present on this MAV, the heading information was obtained from an on-board GPS sensor. For the rotor-craft MAV, the attitude was obtained from a low level complementary filter that combined the on-board gyroscopes, accelerometers and compass data. These measurements along with the resulting localization estimates were transmitted and stored on a ground station for later analysis using a wireless communication network.

The embedded circuit shown in Figure 2.9 was designed for the acoustic processing. This circuit is based on an AVR32 microcontroller that amplifies and digitizes the microphone signals and computes the TDOA localization steps. A sampling rate of 40kHz and a sample size of 1024 was used for the coherence measurements. Furthermore, for computations of the particle filter, a second identical microcontroller was used in parallel. The following sections explain the experiments and results of the three proposed methods. Videos of experiments for each method can be found at http://lis.epfl.ch/ABSMAV.

2.4.1 Localizing the proposed emergency source

Multiple experiments were performed in order to test the performance of the localization system, explained in Section 2.3.1, in locating the proposed emergency sound source while only relying on TDOA measurements. The quadrotor MAV was used to estimate the direction to this sound source from different locations relying only on the perceived sound waves. In these experiments the GPS positions of the target and the MAV were stored and only used as the
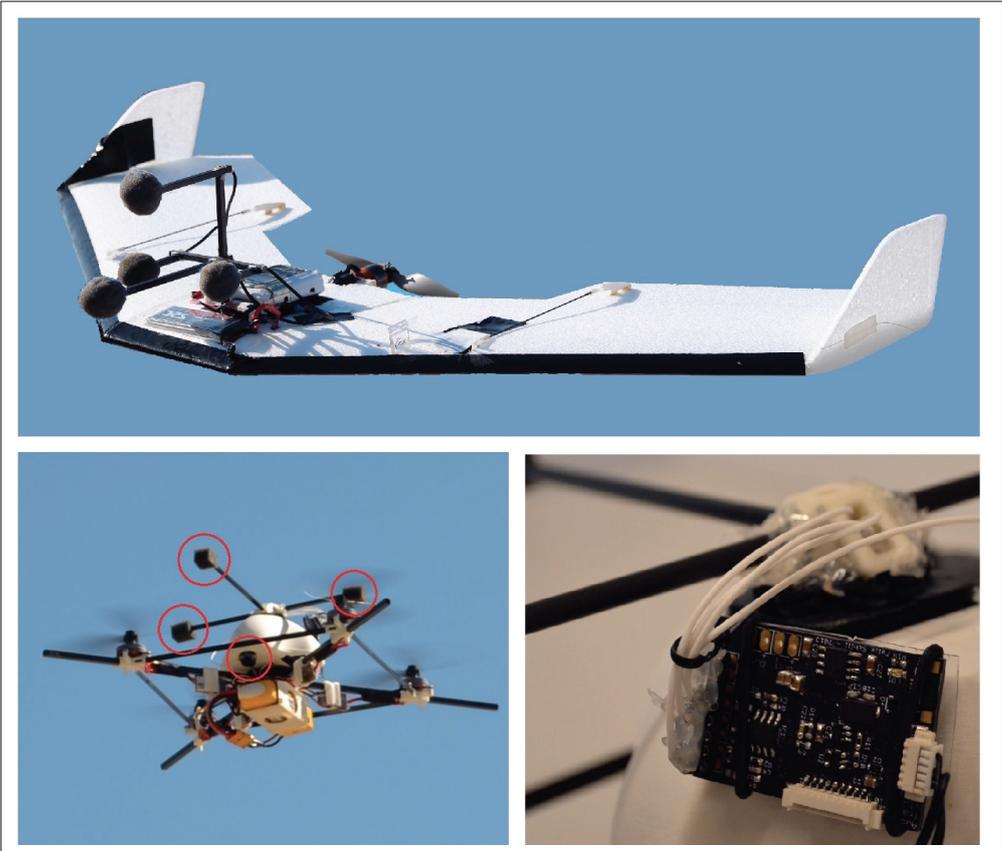
Figure 2.9: Pictures of the two MAV platforms used in the experiments along with the picture of the developed embedded acoustic board for on-board acoustic processing. A microphone array of four microphones is used on each robot. The positions of microphones on the quadrotor is highlighted with red circles
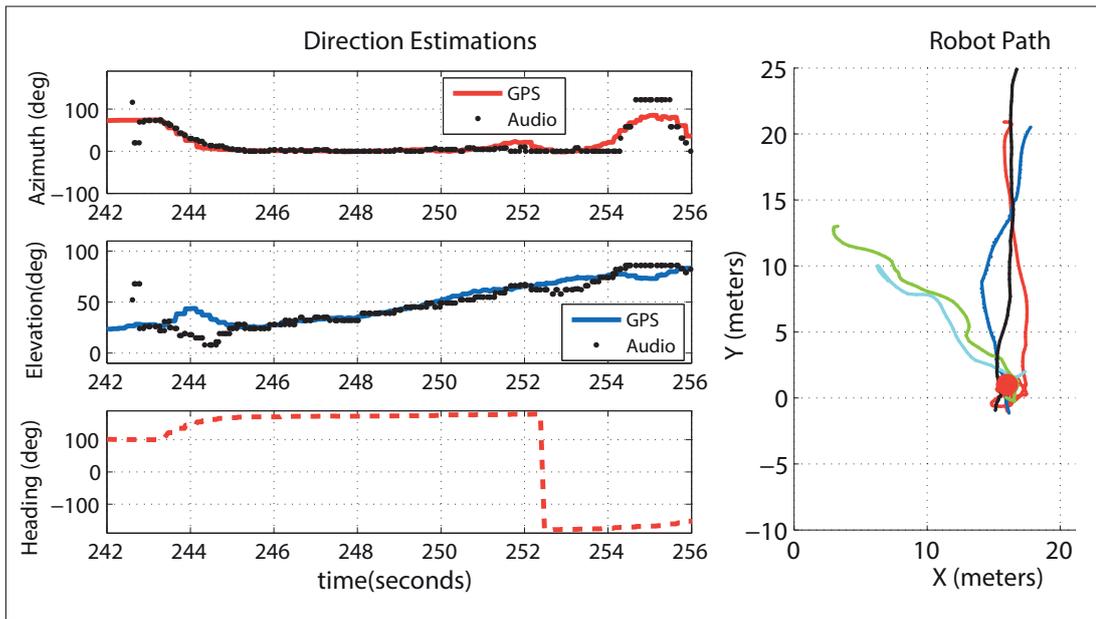
Figure 2.10: Results of the experiments where the rotorcraft MAV was used to locate the proposed acoustic target on the ground by estimating the relative direction of the incoming sound waves and navigating autonomously to reach the target. Left) plots of the audio-based direction estimations and the heading of the robot provided by the IMU. Right) path of the robot and the position of the target for multiple trials.

reference. Experiments showed a good correspondence between the GPS based and the audio-based direction estimates. Figure 2.10 shows the audio-based and the GPS based estimates in one of the experiments. In addition, since with this continuously sound emitting source a continuous direction estimation is obtained, a motion controller was also implemented on the MAV to navigate it to the position of the target. Note that no GPS information is required for this navigation as the robot could reach the target by controlling its attitude and speed to move towards the estimated direction of the target. Experiments showed that the motion controller was capable of successfully navigating the robot towards the target position, by relying on the audio-based direction estimates. The path of the MAV for multiple experiments is shown in Figure 2.10. Note that the motion of the robots could also be observed from the direction estimates, where the initial reduction in the azimuth is due to the robot changing its heading to face the target, the gradual increase of the elevation is because the robot is approaching the target and the elevation of 90° shows that the robot has reached above the target.

2.4.2 Localization using motion exploitation

Experiments were conducted using the fixed wing MAV, for testing the method presented in Section 2.3.2.1, in localizing a commonly available emergency source located on the ground that is emitting a narrowband sound. The MAV was controlled to fly within the visual range

Figure 2.11: (a)The safety whistle used in the experiment (commonly available in outdoor shops). It features two chambers, and emits two closely spaced, superimposed frequencies between 2kHz and 2.1 kHz. (b) A hand-held piezo alarm that emits a single frequency of 3.8 kHz.

of a safety pilot while occasionally reducing or even turning off its engine to increase the detection range by increasing the signal to noise ratio. This reduction in the engine power is achieved automatically whenever the MAV is descending in the altitude. Two different sound sources were used for these experiments: a commonly available safety whistle of 2.1kHz and a 3.8kHz hand-held piezo alarm.The sounds were triggered by a human experimenter in a known location. The whistle was blown in intervals of approximately 1 to 2 seconds.

Figure 2.12 shows the result of an experiment where the safety whistle is used as the target sound source. This figure illustrates the relative direction estimates from the flying MAV compared against the relative direction computed from the GPS positions. The TDOA measurements, dispersion of particles and the direction error between GPS and relative direction is also indicated in this figure. Furthermore the estimated direction is used to obtain an estimate of the target position from triangulation using the MAVs altitude and position. The target position estimation error in meters during a time interval of the flight is also shown in this figure. Here, the particle filter algorithm is initialized after the MAV's motor input drops below a predefined threshold. It can be seen that as soon as the first whistle is observed, particles converge toward the correct direction of the target. Furthermore, when there are no observations available or the MAV's motor input is above a predefined threshold, the particle filtering update step is no longer performed and hence only the probabilistic motion model of the MAV is responsible for the tracking. This results in the gradual increase in the spreading of the particles until the next observation is available. It can be seen that in this experiment, and after the last set of whistle sound observations, the target position is tracked correctly for several seconds and then suddenly the error starts to increase. This is because at this point the MAV starts to perform a sharp 180 degrees turn and the simple forward motion model used does not fully capture this behaviour. Note that although the microphone pair's inter distance and the frequency of the target does not satisfy Equation (2.5), and hence 2 peaks could appear in the cross correlations, there are no ambiguities in the final TDOA observations. This is because the current geometry of the microphone array ensures that for this target frequency not all pairs experience this ambiguity simultaneously. Therefore in the proposed method, in the step in which corresponding time delays among different pairs are identified,

incorrect peaks are automatically eliminated and a single direction measurement is obtained. Hence, in this experiment, as there were no ambiguities present in the TDOA measurement, the Doppler based relative speed estimation step and knowledge of the target frequency were not necessary.

To further test the performance of the system in situations where ambiguities occur in the final TDOA observations, a set of experiments were performed in which a piezo alarm with a frequency of 3.8 kHz was used as the target source. Figure 2.13 shows the result of an experiment that illustrates the correct direction estimation despite the high ambiguity existing in the TDOA measurements. The raw azimuth estimates from TDOA alone are shown as small red dots,indicating ambiguities in the form of 5-6 possible angles are clearly visible. However, the particle-filtered estimate quickly converges to the correct estimate. Furthermore, the target position is shown to be tracked correctly long after observations are no longer available. A detection range beyond 150 meters was achieved with both the whistle and the alarm.

2.4.3 Behaviour based localization

To test the behaviour based approach the quadrotor shown in Figure 2.9 was used in multiple experiments for detecting the direction of a narrowband sound source. Initially, experiments
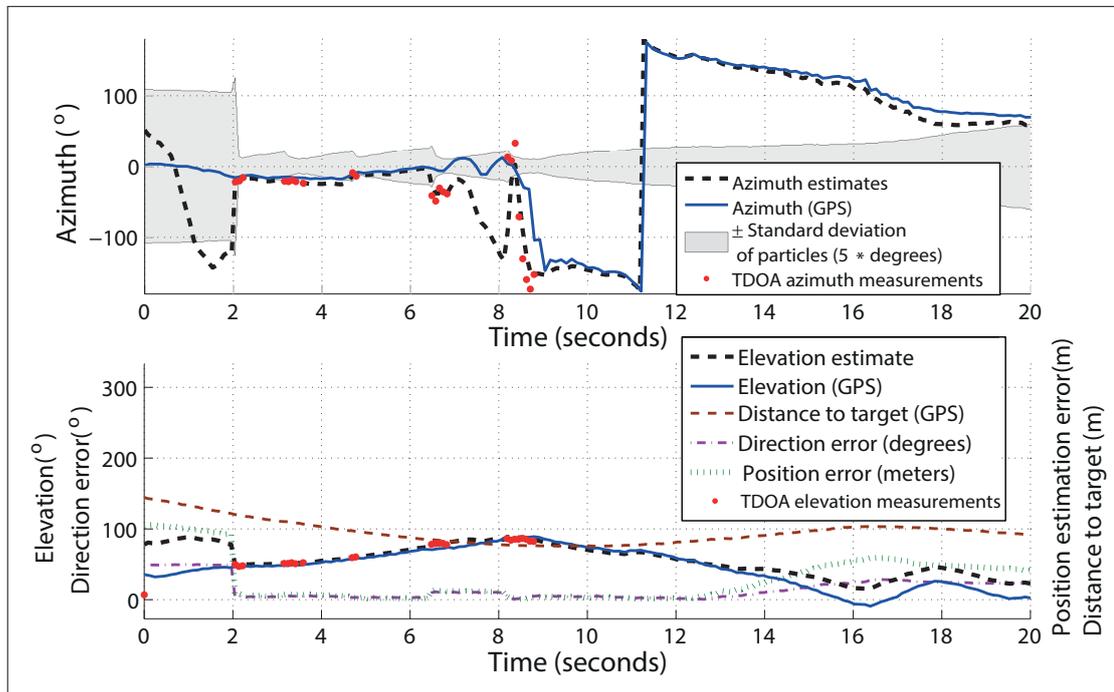
Figure 2.12: Result of an experiment where a fixed wing MAV is used to locate a human target on the ground who is occasionally blowing into a safety whistle. The TDOA measurements also indicate the times that the sound of the whistle is perceived by the MAV illustrating the discontinuities with the sound of the whistle.
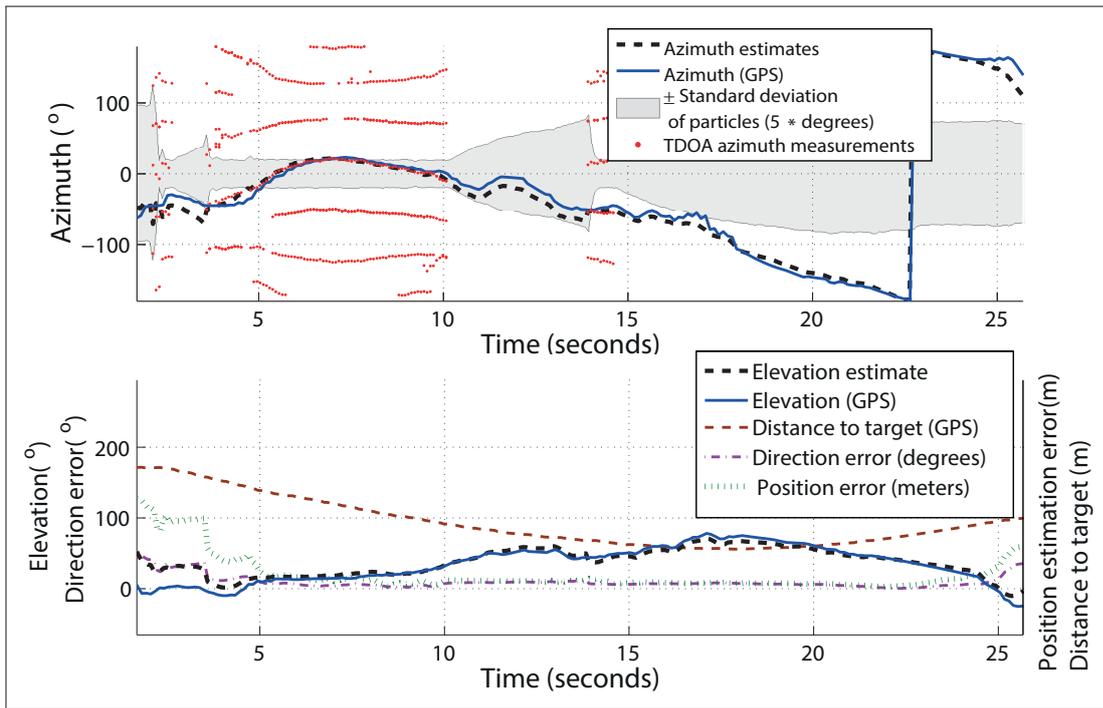
Figure 2.13: Result of an experiment where a fixed wing MAV is used to locate the position of a human target on the ground that is holding a piezo alarm (N=100).

were performed indoors with an 8 camera Vicon system used as the reference. In these sets of experiments, the MAV's heading was changed manually once a sound was played through a speaker. The on-board audio-based relative direction estimates $\vec{u}$ along with the reliability values X, defined in Equation (2.36), were logged throughout the experiments. The MAV's attitude and position along with the position of the speaker were all obtained through the Vicon system and used for computing the true relative source direction. Experiments showed that an initial change in the MAV's attitude once the sound is played always leads to accurate estimations within few seconds whereas keeping the attitude constant always leads to wrong estimates. This confirmed the initial hypothesis that the attitude control by the robot could allow localization of narrowband sources. Figure 2.14 shows the result of two experiments for estimating the direction of a 4.5kHz sound source, where in one experiment the robot's attitude is fixed whereas in the other the attitude is changed. It can be seen that the particles fully converge to the correct direction after a change in the attitude of approximately 140 degrees in the heading.i,e. yaw direction. Experiments also showed that after the initial convergence, the estimator is capable of tracking the target direction with a high accuracy and without the need of further changing its attitude. A root-mean-square-error (RMS) of 2.39 degrees were calculated for the target localization error after the initial convergence of the estimator. In addition, the estimator was robust against discontinuities in the sound play.

To further test the proposed method a set of experiments were performed outdoors for localizing a person blowing into a safety whistle. This time the change in the attitude required for
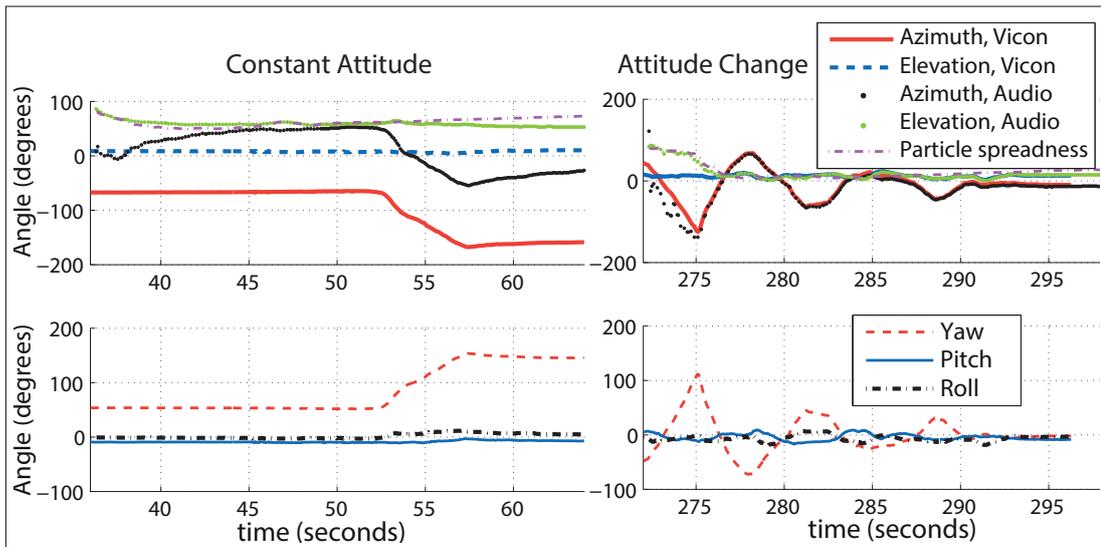
Figure 2.14: Direction estimation in indoor environment, with the behaviour based method, for the two cases of fixed attitude and alternating attitude. Plots start immediately after a sound of 4.5kHz is played. The top plots show the audio-based direction estimations against the true directions computed from the Vicon system. The bottom plots show the robot's true attitude given from the Vicon system. A measure of particle spreadness Ψ is computed from $\Psi = 150\left(1 - \frac{X}{N_P}\right)$, where X is the reliability value in Equation (2.36) to have values in the range $(0, 150)$

obtaining the true direction estimate was performed automatically by the robot. The robot was programmed to maintain its position in the air until the detection of the whistle. This detection was achieved by detecting the occurrence of a strong peak in the sound's frequency spectrum. Upon first detection, the robot started modifying its attitude by firstly increasing and furthermore decreasing its heading by $\approx 140^o$ in approximately four seconds. This was to allow the estimator more time to handle the discontinuities with the whistle sound. Whenever a good convergence in the estimator was detected the robot controlled its orientation to face this estimated direction. A 4.1kHz low cost whistle was blown occasionally by a person in known locations, each time blown for three times with intervals of 1-2 second. Figure 2.15 illustrates four instances of these experiments showing the success of obtaining the correct direction of the target. Plots show clearly the initial change in the attitude, that starts immediately after the sound detection, and the final converged estimated azimuth that was used as the input to the attitude controller to control the final heading of the robot towards the target.
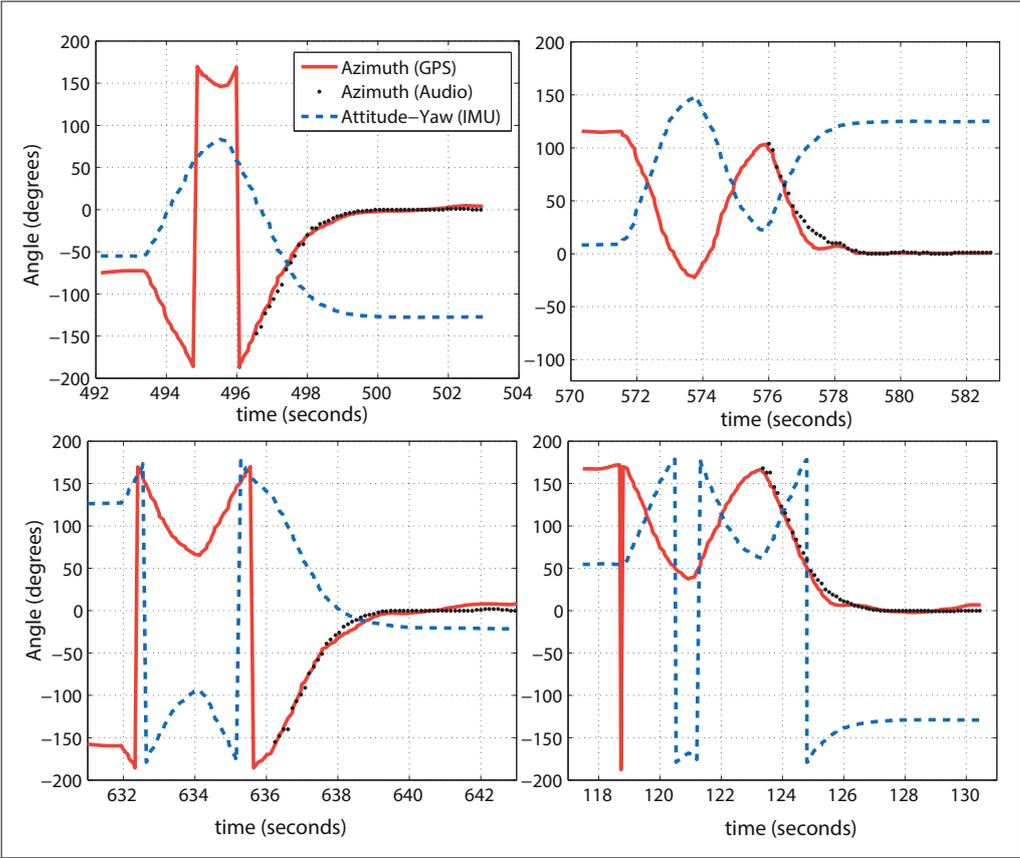
Figure 2.15: Results of experiments where the rotorcraft MAV of Figure 2.9 uses the behaviour based approach to detect a person blowing into a safety whistle of 4.1kHz. The robot was programmed to firstly increase and then decrease its orientation by ≈ 140 degrees, upon detecting the whistle, until a good direction estimation is obtained. Then the robot automatically turns to face the estimated direction.

40

2.5 Conclusion

In this chapter multiple solutions to the problem of localizing emergency acoustic sources on the ground using micro aerial vehicles were proposed. The solutions provided required an on-board microphone array to measure the probable directions of the emergency source based on TDOA measurements, and on-board sensors to obtain information about the state of the MAV. The first solution consisted of designing a sound source that allowed immediate localization by the MAV using entirely acoustic information. The other two methods addressed the problem of localizing the currently available emergency sources that lead to ambiguous TDOA measurements due to the repetitive nature of narrow band sounds. In the second proposed method, knowledge of the vehicle dynamics, along with the sound frequency and the Doppler shift of this frequency were used for resolving these ambiguities. In the third proposed method, autonomous control of the attitude by the robot and furthermore fusion of acoustics and attitude measurements were employed to obtain correct estimations. Investigating different types of MAV motions for improving the localization accuracy and reducing the convergence time, employing multiple MAVs for facilitating area exploration and for achieving cooperative target localization are some potential future directions on this topic.

3 Audio-based relative positioning for multiple micro aerial vehicle systems

DESIGNING a group of autonomous MAVs requires addressing new challenges, such as inter-robot collision avoidance and formation control, where individual's knowledge about the relative location of their local group members is essential. A relative positioning system for an MAV needs to satisfy severe constraints in terms of size, weight, processing power, power consumption, three-dimensional coverage and price. These constraints prevent the current relative positioning systems designed for ground robots and large aerial vehicles to be used in MAVs. Inspired by the sense of hearing in animals [Farnsworth, 2005, Muller and Robert, 2001], which provides them with the ability of using sound for communication and localization, we propose an onboard audio-based system for allowing MAVs in an MAV swarm to obtain information about the position of their neighbouring robots. In this chapter, we firstly propose a method, based on measuring the coherence among signals of a small onboard microphone array, to measure the relative direction of other robots from perceiving their engine sounds in the absence of self-engine noises. We then extend the method to obtain this information in the presence of self-engine noises, for achieving a longer detection range, and for distinguishing the identity of different robots. For this purpose, we propose active acoustic signalling where individuals generate unique chirp sounds similar to birds. A method based on fractional Fourier transform (FrFT) is used by the individuals to identify and extract sounds of simultaneous chirping robots in the neighbourhood. Furthermore, we describe an estimator based on particle filters that fuses the relative bearing measurements with information about the motion of the robots, provided by their onboard sensors, to also obtain an estimate about the relative range of the target robots.

3.1   Introduction

Relative positioning is the problem of gaining information about the position of other robots, by the individuals in a robotics group. In Sections 1.1 and 1.2.2, the importance of an onboard relative positioning system for multi-robot systems was described. It was further described that due to the strict requirements of micro aerial vehicles, requiring a small, compact, lightweight and three-dimensional solution, there is a lack of technological possibilities that could provide this information without relying on any external systems. In this chapter, we propose an onboard audio-based system for individual MAVs that exploits the locally perceived sound waves for obtaining this information.

Audio-based relative positioning has not been favoured for ground robots due to the availability of many other sensor technologies that could be used on the less constrained ground robots and because of the existing challenges in sound source localization inside reverberant and noisy domestic environments. In the case of underwater robotic swarms, the effectiveness of audio-based relative positioning compared to other methods have been shown by some researchers [Kottege and Zimmer, 2007]. In these systems, a pair of hydrophone sensors onboard a small submarine is used for measuring the relative bearing of other sound emitting submarines. Audio-based relative positioning for miniature aerial robots has not been addressed so far. However, existing examples in nature shows the potential success of such a system for aerial robots. Flight calls of nocturnal migratory birds during coordinated migration at night [Farnsworth, 2005], and phonotaxis behaviour among insect swarms for mating and predator avoidance [Gibson et al., 2010, Muller and Robert, 2001] are some of the many existing examples.

An audio-based relative positioning system for swarm of MAVs will have many potential advantages. First of all, this system will be dependant on extremely low cost, lightweight, small size and passive microphone sensors which are very suitable for employment on small scale micro aerial vehicles. Today, microphone sensors with a size of few millimetres are commercially available at a price of around one Euro (as it was illustrated in Figure 1.4). Another important feature of microphone sensors are their omni-directionality, that allows them to provide a full three-dimensional coverage that is an essential requirement for aerial robots as they are operating in the 3D space. In addition, the passivity of these sensors will result in a low power consuming system for having a longer swarm endurance. Design of other acoustic sensors for potential future use on aerial robots have also been investigated in some recent works [Ruffier et al., 2011] [de Bree et al., 2010].

An audio-based relative positioning system will be based on sound waves that are independent of illumination and weather conditions, such as fog, dust, rain and smoke, and can operate at night time. This system does not necessarily need a direct line-of-sight between the robots for its operation as sound waves are capable of overcoming obstacles throughout their ways, hence it can provide information through foliage and occlusions caused by other robots. In addition, such a system could potentially be less computationally expensive compared to

vision-based systems, as it will mainly rely on the phase information that is already available in the one dimensional sound waves, rather than the need for feature extraction and processing sequences of two-dimensional high-resolution images to find and locate these small robots. Furthermore, as the engine of most flying robots already produce sound waves while flying, this sound could be exploited for obtaining passively the relative position of team-mates and other non cooperative flying platforms. Finally this system could be exploited further to detect, locate and study many other interesting acoustic targets from the flying MAVs which introduce many potential applications, such as locating emergency acoustic sources in search and rescue missions that was described in Chapter 2.

This chapter is organized as follows: Section 3.2 describes the relative bearing measurement system for passive localization of neighbouring robots based entirely on the sound of their engines, with the related experiments and results presented in Subsection 3.2.5. In Section 3.3 active acoustic signalling is introduced, to resolve some of the limitations of the previous method, where individuals generate bird-like sounds to assist others in locating them. Furthermore, in this section an estimator based on particle filters is developed for estimating the relative locations of other robots throughout time. Experiments and results related to this section are presented in Subsection 3.3.4. Finally, Section 3.4 concludes the chapter by providing a conclusion and some potential future works on this topic.

3.2 Passive audio-based relative-bearing measurement system

This section explains our work on obtaining an onboard system for measuring the relative bearing to neighbouring robots based entirely on the sound of their engines. Exploiting the already available sound produced by the engine of flying robots would result in a high energy-efficient solution and could potentially be employed to locate many other non-cooperative aerial platforms, that is an important feature for having a reactive sense and avoid system to avoid mid-air collisions. The proposed method is suitable for the cases when there are either no self-engine noise present, or the engine sound from other robots and the self-engine noise have the required differentness for separating the sounds from the self-engine noise. This method could potentially be employed in groups of fixed-wing MAV's as they are capable of gliding with their engines reduced or turned off, and in groups of rotor-craft MAVs that can rest on the ground or attach to ceilings or walls [Stirling et al., 2012, Mellinger et al., 2010, Doyle et al., 2013, Roberts et al., 2008]. Localizing the engine sound of distant robots in presence of an engine with the same sound characteristics that is only few centimetres away from the microphones, hence resulting in an extremely low signal to noise ratio (SNR), is a challenging problem that is beyond the scope of this work. However, a method based on active acoustic signalling is described later in Section 3.3 to obtain the relative positioning in presence of the self-engine noises.

The proposed system is based on the TDOA sound source localization method that was previously introduced in Section 2.2, and a compact on-board microphone array used for
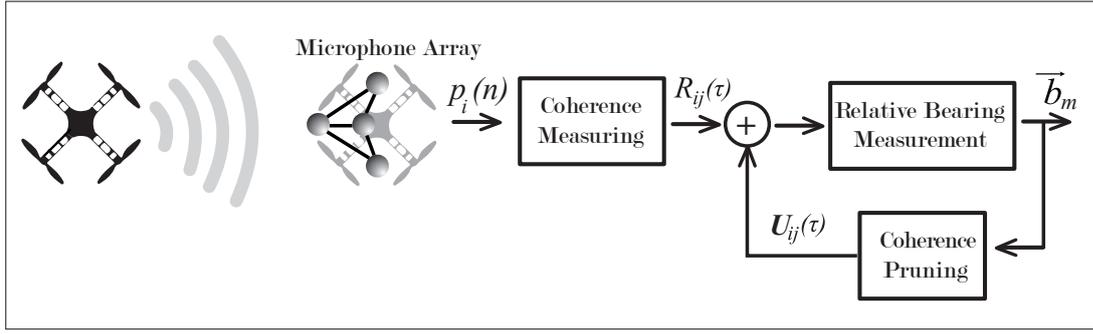
Figure 3.1: Schematic diagram of the proposed passive audio-based relative bearing measurement system illustrating the main parts of the system.

measuring the sound waves. Note that unlike the emergency sound sources that had a narrowband nature, the engine sound of robots and other flying platforms are mostly broadband sounds and consist of a wide range of frequencies. Hence the problem of ambiguous direction measurements, described for the narrowband emergency sources is not experienced here. Figure 3.1 presents the block diagram of the passive relative bearing measurement system consisting of its main units. An explanation of each individual unit is provided in this section. Some of these units will also be the building blocks of the other proposed methods and will be used in the next sections.

3.2.1 Microphone array

An onboard microphone array of four microphones is used for simultaneously measuring the sound waves, emitted from the engine of other robots at four different locations. Four microphones are chosen in this work since it is the minimum number of microphones required, if they are not all placed on the same plane, to locate the direction of sound sources in a 3D space without ambiguity. This minimum number is chosen in order to minimize the hardware and computational complexity, for achieving a simple and real-time solution suitable for small scale MAVs. However, due to the small size of these sensors, more microphones could easily be added on the MAVs to increase the robustness of the system. There are no strict constraints on the position or the geometry of the microphone array, that makes it suitable for being mounted on any type of MAVs, and only the microphone positions relative to each other must be known with a good accuracy. In this section, we use two different type of microphone array geometries, as shown in Figure 3.2, where a flat triangular microphone array, with a total weight of 3.5 grams, is used on pocket-size rotor-craft MAVs and a regular tetrahedral microphone array is used on the fixed-wing MAVs. The regular tetrahedral microphone array geometry provides approximately equal localization performance in all directions [Hu et al., 2011]. The flat microphone array has the advantage of compactness, but it introduces a top-bottom ambiguity when localizing sound sources in the 3D space as all the microphones are placed on the same plane. However, because the localization with this array is performed after the rotorcraft MAV rests on a flat surface, this ambiguity is not experienced and the robot is
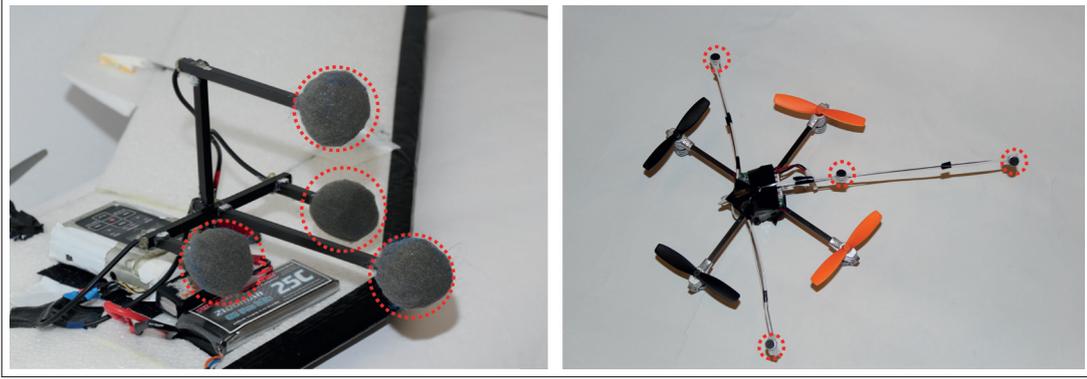
Figure 3.2: Pictures of the two different microphone arrays used in this section, with the microphones indicated by red circles. Left: a tetrahedral microphone array with inter-microphone distance of 10cm is used on the nose of a fixed wing MAV platform. Microphones are covered by foam for wind-protection. Right: a flat microphone array attached on a pocket size rotorcraft platform, with three microphones forming an equilateral triangle with sides of 18cm and the fourth microphone in the center of this triangle. The total weight of this array is 3.5 grams.

aware that sound sources are located only on one side of the array.

Initially, sound waves are picked up by the microphones and are converted into electric signals. These signals are then amplified and filtered using basic analogue filters to remove the unwanted electrical noises. Furthermore, signals from all the microphones are simultaneously sampled and converted into a digitized form with a sampling frequency of F_s. Upon obtaining a discrete sequence of N samples from all the microphones, these sequences are passed to the coherence measuring unit. The length N of the sequence should be chosen as a trade-off between the the stability, obtained with large values of N, and the tracking requirements, if there are no limitations in the available memory and the processing power. However, in this work, we use $N = 1024$ samples to meet the low memory and computational power of common micro-controllers.

3.2.2 Coherence measuring unit

The coherence measuring unit is based on the method that was previously explained in Section 2.2.1, where cross-correlation is used to obtain a measure of similarity between all of the microphone signals. The unit starts by measuring the similarity between each pair of microphone signal sequences, as a function of time delay τ applied to one of the sequences, using the function:

$$R_{ij}(\tau) = FFT^{-1}\left(\frac{FFT\left(p_i(n)\right).FFT^*\left(p_j(n)\right)}{W}\right) \tag{3.1}$$

where FFT is the Fast Fourier Transform, FFT^{-1} is the inverse FFT, FFT^* denotes the complex conjugate of the FFT results, $p_i(n)$ is the digitized sequence obtained from microphone i with $n = 0, 1, ..., N-1$ and τ is the correlation lag in samples. W is the spectral weighting function for improving the similarity analysis which is described later in this section. Note that for $W = 1$ the coherence measuring function (3.1) is a general cross correlation that is computed in the frequency domain by taking the inverse Fourier transform of the cross spectrum. The value of $R_{ij}(\tau)$ for all the probable discrete time delays in the range $\pm \tau_{\max}$ is only of interest and is stored for further processing. The maximum time delay $\tau_{\max}$ is limited to the sampling frequency Fs and the distance d_m between the two microphones:

$$\tau_{\max} = \frac{d_m}{c} F_s \qquad (3.2)$$

where c is the speed of sound.

Investigating the similarity degree only for a set of integer delays limits the resolution of the coherence measurements, and hence the bearing measurements, to the sampling frequency and the microphone pair inter-distances. In this work a sampling frequency of $40KHz$ is used, which allows the digitization to be also performed by the same micro-controller that is used for the computations and avoids the need for any additional analogue to digital converting hardware. However, to provide more resolution to the similarity measurements, a cubic spline interpolation with a factor of 10 is performed on the values of $R_{ij}(\tau)$ within the desired range. The sub-sample resolution is computed from:

$$R_{ij}\left(\tau + \frac{l}{10}\right) = R_{ij}(\tau) + \frac{l}{20}\left[R_{ij}(\tau+1) - R_{ij}(\tau-1) + \frac{l}{10}\left(2R_{ij}(\tau-1) - 5R_{ij}(\tau) + 4R_{ij}(\tau+1) - R_{ij}(\tau+2) + \frac{l}{10}\left(3R_{ij}(\tau) - 3R_{ij}(\tau+1) + R_{ij}(\tau+2) - R_{ij}(\tau-1) \right) \right) \right]$$
$$(3.3)$$

where $l = 0, 1, 2, ..., 10$. Figure 3.3 illustrates the result of the coherence measuring, before and after the interpolation, for a pair of microphone that is inside a sound field caused by the engine of a flying rotor-craft MAV. It can be seen that the interpolation leads to a higher time delay resolution and smoother peaks which increases the position accuracy of the maxima points.

One limitation of using the general cross correlation method, i.e. $W = 1$, for measuring the similarity, is that results are strongly dependant on the statistical properties of the sound signal [Valin et al., 2007], and it usually leads to wide cross correlation peaks. The role of the weighting function W, in Equation (3.1), is to improve the localization robustness by performing sound frequency bin weighting. Many different weighting functions are proposed in the literature [Perez-Lorenzo et al., 2012, Miro, 2007]. The PHAT weighting, proposed by [Knapp and Carter, 1976], is one of the most popular weighting functions among research works in the sound-source localization community, showing robust localization performance,
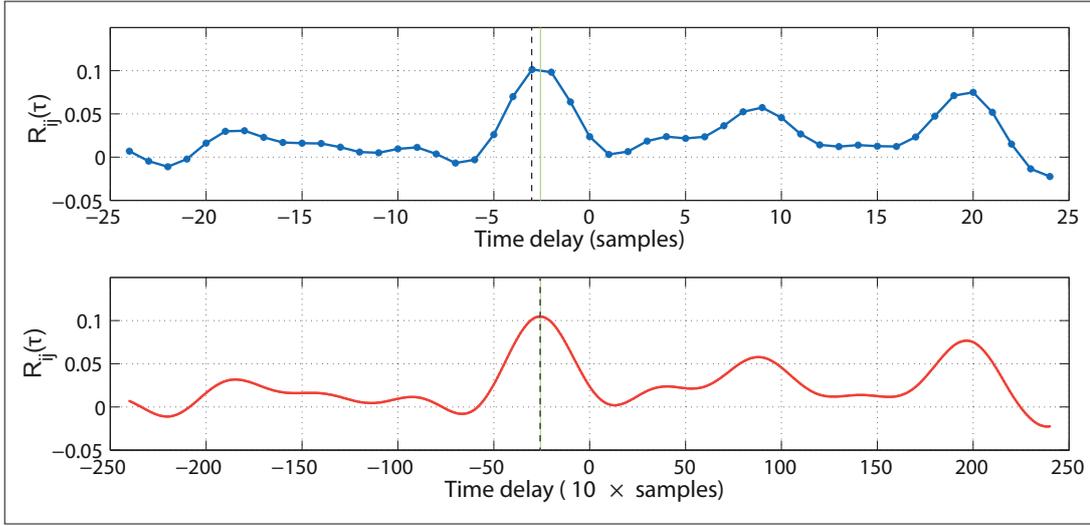
Figure 3.3: Coherence measure $R_{ij}(\tau)$ for a pair of microphones inside a sound field, top: before interpolation, bottom: after interpolation. Dashed vertical line shows the time-lag of the maximum similarity and the solid vertical line shows the true time delay computed from the true source and microphone positions.

particularly against reverberations:

$$W_{PHAT} = \left| \text{FFT}\left(p_i(n)\right) \right| \left| \text{FFT}\left(p_j(n)\right) \right| \tag{3.4}$$

This weighting function whitens the cross-spectrum in order to give equal contribution to all of the frequencies. This way the coherence measurement is only based on the phase information of the signals, and it is independent of the signal dynamics, resulting in a much sharper correlation peaks that could increase the precision. One drawback of this method is that every frequency bin in the spectrum contributes equivalently in the similarity measure, even if the signal is not present in that frequency bin or it is dominated by noise. A modified version of the PHAT weighting was used here instead to only take into account the frequency bandwidth where the sound is mostly present and to also de-emphasize the dominant frequency components within this range.

$$W = \chi \left(\left| \text{FFT}\left(p_i(n)\right) \right| \left| \text{FFT}\left(p_j(n)\right) \right| \right)^{\alpha} \tag{3.5}$$

where

$$\chi = \begin{cases} 1 & f_{\min} < f < f_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

with $f_{\min}$ and $f_{\max}$ being the minimum and maximum limits of the frequency range. The value of α is within the range $[0-1]$, and it is used to control the trade-off between full whitening and no whitening for the frequency bins inside the defined range. The larger the values of α, the sharper are the peaks in the resulting coherence measurements. However, a large α
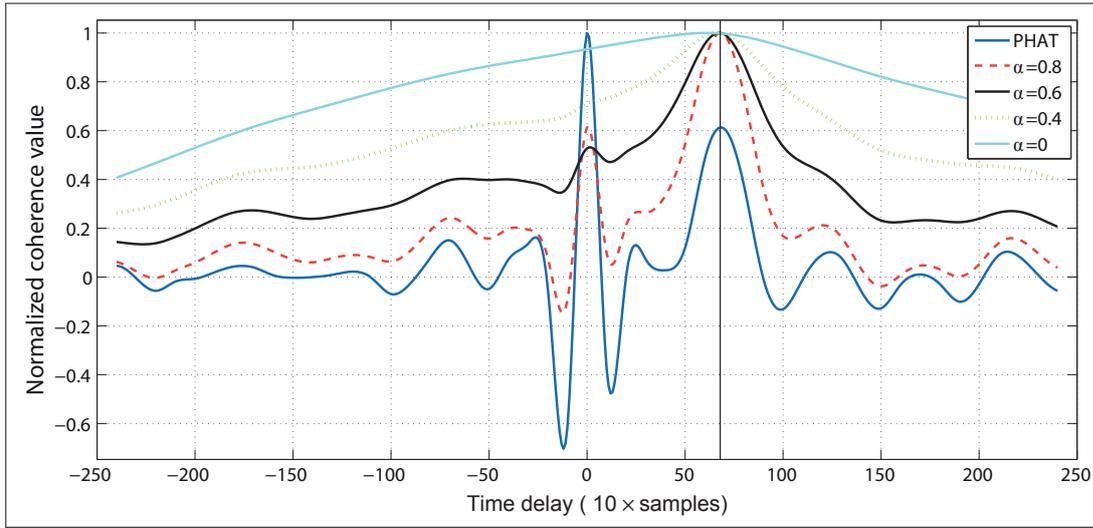
Figure 3.4: Coherence measure $R_{ij}(\tau)$, between signals of a microphone pair that is experiencing the sound of a rotorcraft MAV in an indoor environment, with: PHAT weighting, $\alpha = 0.8$, $\alpha = 0.6$, $\alpha = 0.4$ and general cross correlation, i.e. $\alpha = 0$. The entire frequency range was used for computing the coherence measurements.

could lead to less robustness in case of low signal to noise ratio situations. In addition, a larger frequency range would also result in sharper correlation peaks. Hence, it is desirable to have engine sounds that cover a wide range of frequencies.

Figure 3.4 shows a comparison between, PHAT weighting, $\alpha = 0.8$, $\alpha = 0.6$, $\alpha = 0.4$ and general cross correlation, i.e. $\alpha = 0$, for a pair of microphones experiencing sound of a rotorcraft MAV in an indoor environment, where the entire frequency range was used for these coherence measurements. It can be seen that the PHAT weighting results in an erroneous global peak which does not correspond to the source. This is because the sound of the robot's engine is absent in a wide range of frequencies and the noise present in these frequencies are considered equally in computing the coherence. This problem can be resolved, by reducing the value of α to put more emphasize on the dominant frequency bins, with the expense of obtaining wider peaks.

Figure 3.5 shows an example for the effect of the frequency range on the coherence measurements, where F_{min} is equal to zero and F_{max} is reduced from the maximum possible frequency, i.e. PHAT weighting, to the frequency of 2kHz. It can be observed that the reduction of the frequency range weakens the false peak in the PHAT weighting. This is since the sound produced by the used rotorcraft engines is mostly present in the lower frequency bins. On the other hand, too small frequency range results in wide and inaccurate peaks.

The three parameters of α, $f_{\min}$ and $f_{\max}$ can be tuned according to the sound-source specifications and the available signal to noise ratio, for obtaining the best performance. The instantaneous frequency spectrum of the measured sound sequence could be used for on-line
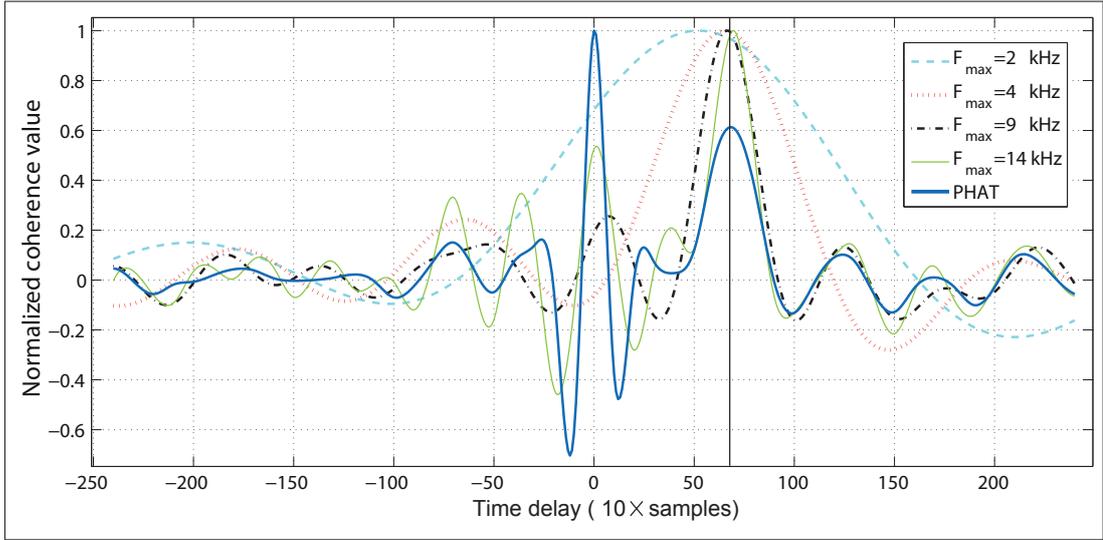
Figure 3.5: Coherence measure $R_{ij}(\tau)$, between signals of a microphone pair that is experiencing the sound of a rotorcraft MAV in an indoor environment, for different frequency ranges and all with $\alpha = 1$.

computation of these parameters prior to measuring the coherence, using a simple thresholding technique. The spectrum of the sound measurement for when there are no target robots present, could be used as a reference for obtaining suitable threshold values. Figure 3.6 shows the frequency spectrums for two sound sequences measured when "no robots" and when "a single robot" was present, in the same experimental setup as for Figures 3.4 and 3.5. It can be seen that a suitable frequency range can be obtained from comparing the two spectrums to find the range where the robot sound is mostly present. Also since in the defined range there are very few frequency bins that does not contain the sound of the robot, a value of α close to one would be suitable to consider equally all of the frequencies inside this range .

The instantaneous frequency spectrum of the measured sequences could also be employed to detect the presence of robots in the vicinity and to distinguish the sound of a robot from other sound sources that might be present. For this, a template of the robot sound's frequency spectrum is stored in the memory and is used to compute the similarity of incoming sounds with this template, prior to the coherence measurements. A similarity value S is computed using the cross correlation method:

$$S = FFT^{-1}\left(P_{mic}(k)P_{tmp}(k)^*\right) \qquad k = 0, 1, ..., N-1 \qquad (3.7)$$

where P_{mic} is the frequency spectrum of one of the microphones and $P_{tmp}(k)$ is the frequency spectrum of the stored template. A good similarity value, defined by a threshold that is found empirically, indicates the sound belongs to the desired robots.
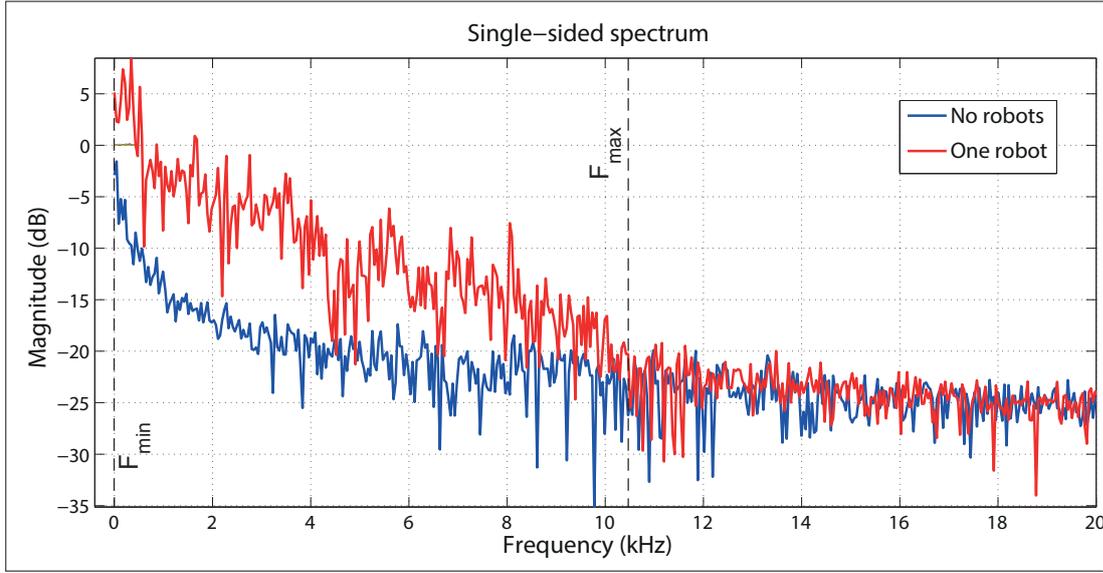
Figure 3.6: Frequency spectrums of a sound sequence in two cases of "no MAVs" and "a single MAV" being present in proximity of a microphone. This information could be used to find good parameters for the proposed weighting function

3.2.3   Relative bearing measurement unit

Upon finding the similarity measure R_{ij} from (3.1) for all microphone pairs ij, a search for the most likely source direction $\overrightarrow{b}_m$ is performed. This is the direction that maximizes the sum of the coherence measurements from all of the microphone pairs.

$$\overrightarrow{b}_m = \arg\max_{\overrightarrow{b}} \sum_{i,j} R_{ij}(\tau_{\tilde{b}ij}) \tag{3.8}$$

where time delay $\tau_{\tilde{b}ij}$ is the expected time delay if the source was in the direction $\overrightarrow{b}$, and is computed from the coordinates of microphones i and j in the body fixed coordinate system.

$$\tau_{\tilde{b}ij} = \frac{\overrightarrow{x}_{ij} \cdot \overrightarrow{b}}{c} \qquad \overrightarrow{b} = (b_x, b_y, b_z)$$
$$\overrightarrow{x}_{ij} = (m_{ix} - m_{jx}, m_{iy} - m_{jy}, m_{iz} - m_{jz}) \tag{3.9}$$

where (m_{ix}, m_{iy}, m_{iz}) are the coordinates of microphone i and c is the speed of sound.

A spherical geodesic grid, defined on a unit sphere, is used to search for the most likely source direction among the set of all potential directions. Each grid point represents a direction vector $\overrightarrow{b}$ that starts at the origin and ends at that grid point. A geodesic grid of 2562 points is used in this work to cover the entire 3D directions (see Figure 3.7). For the case of a resting rotor craft, this is reduced to 1313 points to only cover the potential directions described by a half sphere over the resting surface. For having a fast searching speed that is suitable for
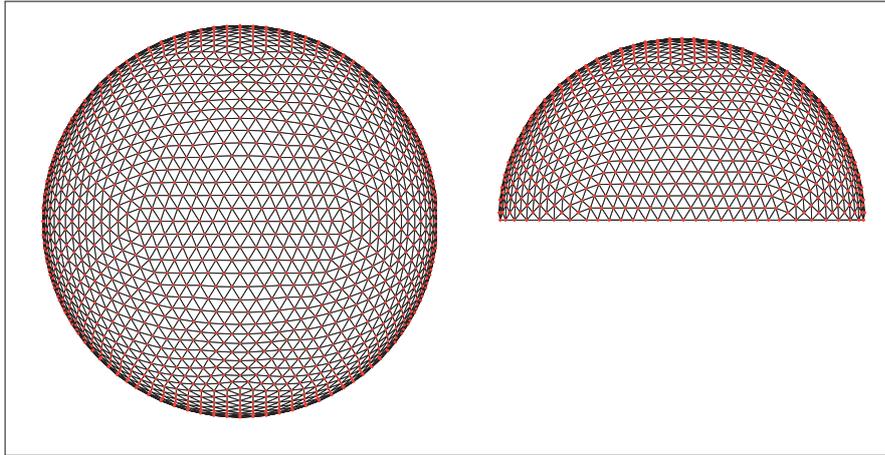
Figure 3.7: The grid used in the bearing measurement unit for finding the most probable sound direction. Left: A geodesic grid of 2562 points to represent the set of all 3D directions. Right: A geodesic grid of 1313 points is used to represent potential directions for a robot resting on a surface.

the real-time implementation, the time delays $\tau_{\vec{b}ij}$ for all the grid points are computed in advance and stored in a lookup table in the static memory. The bearing measurement unit then simply goes through the table and uses the stored time delays to compute $\sum\limits_{i,j} R_{ij}(\tau_{\vec{b}ij})$ for all the directions and then finds the direction with the maximum value. Figure 3.8 shows the result of a grid search from an experiment involving two small rotorcraft MAVs, one resting and one flying MAV, which illustrates the likelihood of all the potential relative directions.

To improve the resolution of the direction estimation, that is limited by the resolution of the grid, a weighted averaging between the grid point having the highest coherence value and its six adjacent points, based on their coherence values, is used as the final estimated target direction.

3.2.4   Coherence pruning unit

The previous units described a method for locating the direction of a single neighbouring robot by perceiving the sound of its engine and finding the direction with the maximum coherence among all of the microphone-pair signals. In the case of multiple neighbouring robots, this method will provide the direction of the dominant sound source that is exhibiting the highest coherence in the similarity measurements. For a homogeneous MAV group with equivalent engines, that are producing the same sound characteristics, the dominant sound source will correspond to the nearest neighbouring robot. However, in practice, the dominant sound might not always be the nearest neighbour as the engine sounds are varying with other factors such as the robot's speed and throttle power. The Coherence pruning unit is used to also obtain the bearing information of other existing robots that are potentially masked behind the sound of the dominant robot. Inspired by the work of [Brutti et al., 2008b], the idea
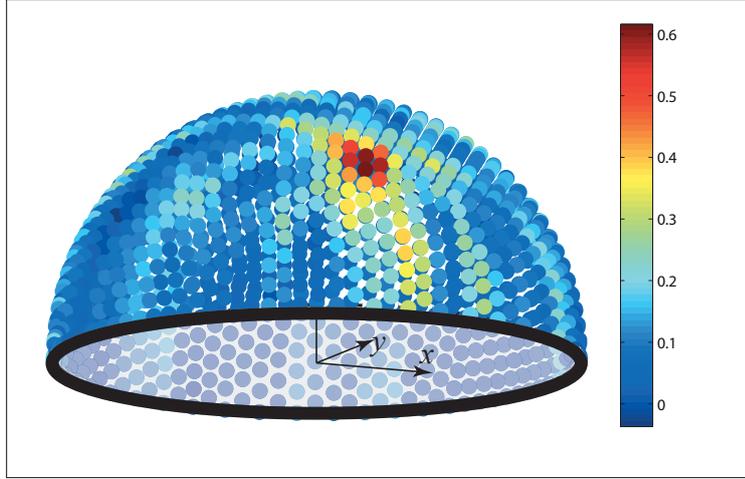
Figure 3.8: Grid search result for an experiment involving a resting rotorcraft MAV and a target flying rotorcraft MAV, showing the likelihood of the target's 3D direction for all potential directions around the resting MAV. Each cell is on the surface of a unit sphere and represents the end of a unit vector starting from the origin. The cell with the highest value is considered as the direction of the target robot

is to de-emphasize the effect of the dominant robot in all of the similarity measurements in order to locate other robots with weaker coherences.

The coherence pruning unit initially uses Equation (3.9) to compute the time delay $\tau_{\vec{b}_m ij}$, for every microphone pair, that corresponds to the resulting bearing measurement $\overrightarrow{b}_m$ from the bearing measurement unit. This time delay is then used to generate a pruning sequence $U_{ij}(\tau)$ that is added to the similarity measurements $R_{ij}(\tau)$ to de-emphasize the existing peak of the dominant source at time lag $\tau_{\vec{b}_m ij}$:

$$R_{ij}^o(\tau) = R_{ij}(\tau) + U_{ij}(\tau) \tag{3.10}$$

$$U_{ij}(\tau) = \begin{cases} B(\tau) & B(\tau) > 0 \\ 0 & B(\tau) \leq 0 \end{cases} \tag{3.11}$$

$$B(\tau) = \frac{1}{L}\left(\tau - \tau_{\vec{b}_m ij}\right)^2 - R_{ij}(\tau_{\vec{b}_m ij}) \qquad \tau \in [-\tau_{\max}, \tau_{\max}] \tag{3.12}$$

where $B(\tau)$ is a second order polynomial sequence and L is a constant that defines the sharpness of this polynomial. The value of L is chosen to produce a polynomial with an approximately equal sharpness as the coherence peaks, that is dependant on the parameters used in the similarity measurements that were explained in Section 3.2.2. Figure 3.9 shows plots of a microphone pair's coherence measurements in an experiment with one perceiving MAV and
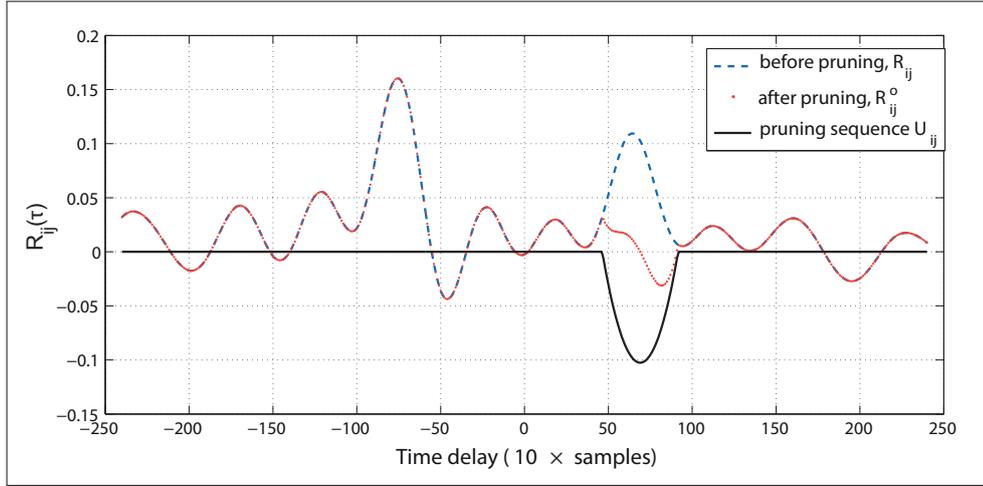
Figure 3.9: Coherence measurement plots in an experiment with one perceiving rotorcraft MAV and two target rotorcraft MAVs, illustrating the pruning step

two target MAVs, before and after the the coherence pruning was preformed. Note that the localized dominant sound source does not necessarily correspond to the dominant peaks in the coherence measurements, as shown in this figure.

The pruned coherence measurements R_{ij}^o are then passed into the bearing measurement unit to search for the most probable sound direction. This procedure is repeated similarly for locating other existing robots that might also be present in the perceived sound mixture. Figure 3.10 shows the direction likelihoods, obtained by the bearing measurement unit, for multiple pruning iterations, on a time window of sound measurements containing the sound of four rotorcraft MAVs.

Experiments, discussed later in Section 3.2.5, shows that up to two pruning iterations can be performed to acquire an accurate direction for the three most dominant target robots. A large drop in the precision after the third pruning iteration, for obtaining the direction of the fourth dominant robot, is observed. However, in reality, since the dominancy of robots alter in time, due to the change in the engine sound and the movements of robots, it is potentially possible to track the direction of more robots throughout time with only two pruning iterations and a memory based algorithm.

The number of pruning iterations should be in accordance with the number of target robots. Always performing two pruning iterations for finding the three dominant targets, regardless of the number of robots that are present, results in false detections for the cases with less than three target robots. If this information is unavailable, the total cross correlation value $\sum_{i,j} R_{ij}(\tau_{\vec{b}_m})$ for the most probable direction $\vec{b}_m$, after every pruning step, and a simple thresholding method, could be used to disregard detections with poor coherence values. An alternative methods for identifying the false detections is to check the direction consistency in few adjacent time frames. Figure 3.11.a shows an example of performing two pruning steps, to
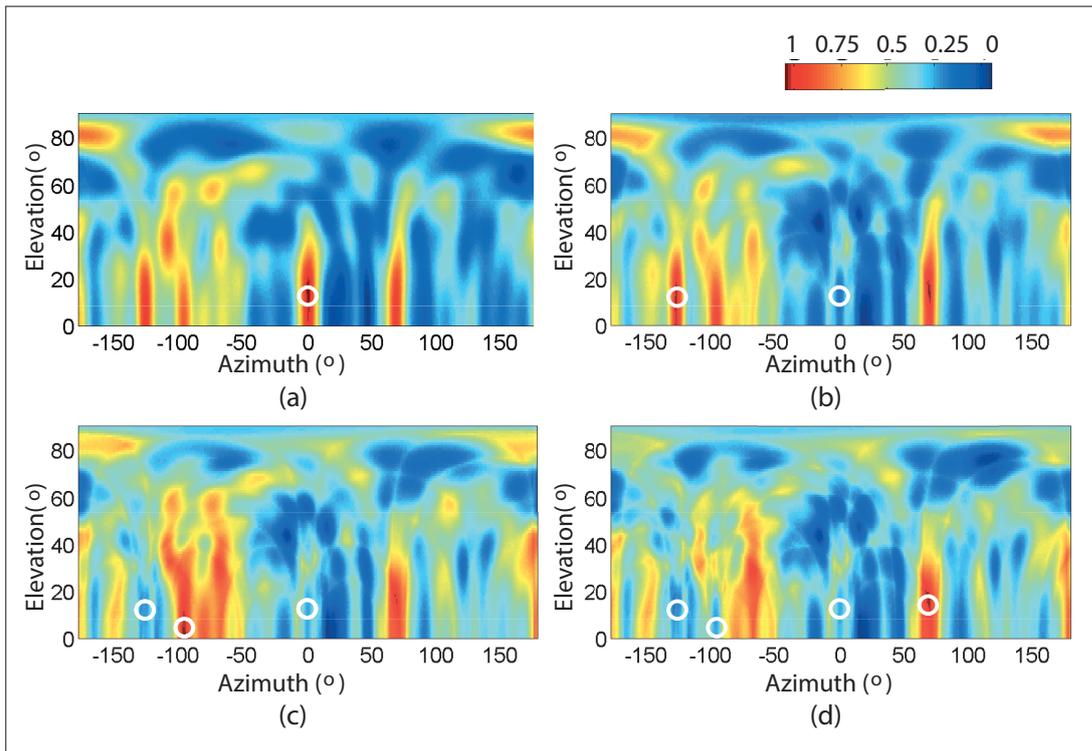
Figure 3.10: The normalized direction likelihood patterns for multiple pruning iterations, using a single time window of sound measurements containing the sound of four rotorcraft MAVs. The direction with the maximum likelihood, i.e. direction with the highest total cross correlation value $\sum R_{ij}(\tau_{\tilde{b}ij})$, after each pruning iteration is considered as a new robot and is marked by a new white circle. a) No pruning is performed, showing the direction of the dominant robot. b)After the first pruning, showing the direction of the second robot c) After the third pruning, indicating the direction of the third robot. d) After the third pruning, showing the direction of the fourth robot. The accuracy of the direction measurements are discussed in Section 3.2.5

locate three target robots, in an experiment where only two target robots were present. A clear inconsistency in the directions, obtained after the second pruning iteration, between multiple consecutive time frames is observed that indicates only two target robots are present in the proximity.

Another alternative, but effective, method for accurately detecting the number of target robots in a single time frame, that could also provide information about the identity of the neighbouring robots, is to exploit the pulse width modulation (PWM) signal of the robots' motor drivers for adding a unique acoustic tone to the engine sound of every robot. The PWM signal is an electric signal provided by a robot's motor driver that controls the motor speed by adjusting the pulse width of this signal. A strong acoustic tone is generated by the motors with the same frequency as the PWM signal. Figure 3.11.b shows the spectrogram of the measured sound waves of a single electric motor equipped with a propeller inside an anechoic chamber,
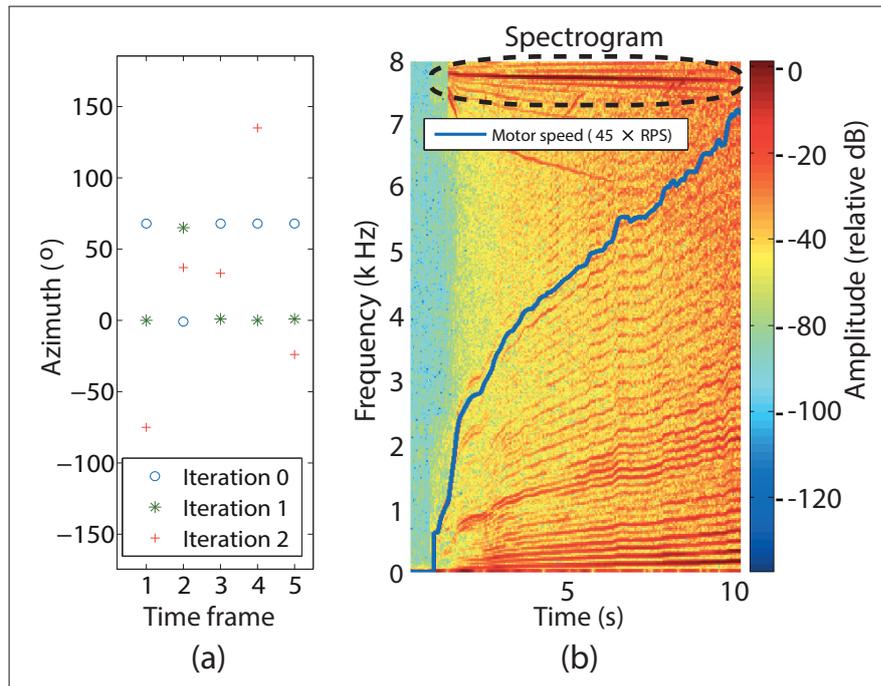
Figure 3.11: a) Direction measurement by performing two pruning steps, to locate three target robots from their engine sounds, for an experiment where only two target robots were present. The inconsistency in the direction measurements obtained after the second pruning iteration, between multiple consecutive time frames, is an indicator that only two robots are present in the detection range. b) Spectrogram of sound recording, in an anechoic chamber, from a single electric motor equipped with a propeller, along with a plot of the motor's rotational speed recorded by a tachometer. The tone generated by the motor itself due to the input 7.8 kHz PWM signal, indicated by the dashed ellipsoid, has a constant frequency that is independent of the motor speed, whereas a direct relationship between the rotational speed of the motor and the sound of the propeller exist.

indicating that the tone due to the PWM signal is independent of the motor speed. It is possible to generate a unique tone for every robot by assigning a different PWM frequency to each robot. This allows the perceiving robots to use their measured sound frequency spectrum, $FFT(p_i(n))$, from one of the microphones, and simply count the number of target robots. Our research on this topic showed that it is also possible to generate more complex acoustic signals, instead of a simple pure tone, by dynamic variation of the PWM frequency without affecting the motor speeds.

3.2.5 Experiments and results

Several experiments were performed to test the proposed audio-based relative bearing measurement system on real robotic platforms. Experiments showed the success of the system in obtaining accurate measurements in the absence of the self-engine noise and in low acoustic
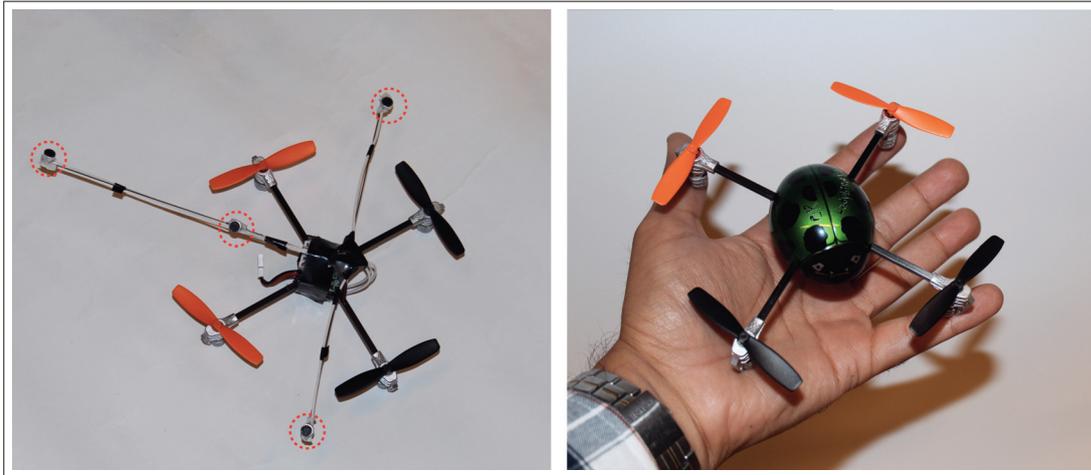
Figure 3.12: Pictures of the pocket size MAVs used in indoor experiments for testing the proposed passive relative bearing measurement system. The perceiving MAV is equipped with a flat microphone array of 4 microphones with a total weight of 3.5 grams. The microphones are marked by the circles and form a triangle of length 18cm with a microphone in the center of the triangle.

noise environments. However, since the detection range of the system is directly dependant on the output sound pressure level (SPL) of the engines and the environmental acoustic noise level, the suitability of the approach highly depends on the type of used MAV platforms and the operating environments. In the case of the platforms used in this work, a good performance were observed in both indoor and outdoor environments. Precise relative bearing measurements were obtained indoors, even with pocket size rotorcraft MAVs having small and quiet engines. Figure 3.12 shows the picture of the two pocket size MAVs that were used for the indoor experiments.

Figure 3.13 shows the result of an experiment in measuring the 3D relative bearing to a pocket-size flying rotor-craft MAV from a perceiving stationary team-mate resting on the ground. The target MAV was flown manually inside an empty room with dimensions $(6 \times 3.5 \times 3)$ meters with the perceiving MAV located in the center of the room. Motion tracking cameras, i.e. Vicon system, was used for measuring the true robot positions and for computing the actual relative bearings between the robots. No other major sources of acoustic noise was present inside the room, however, the sound of the cooling fans belonging to the 8 tracking cameras and two computers could be perceived clearly. Accurate direction estimates with only few outliers were observed throughout the experiments despite the soft sound generated by this small scale robot. Figure 3.14 shows the error histograms of the relative azimuth and elevation errors for this experiment. Due to the existing inconsistency with the azimuth errors in the spherical coordinate representation, that could show large azimuth errors for small angular differences near the poles with elevation of 90°, the shortest angular error between the measured and the true 3D direction vectors was computed for having a better representation of the system's precision. The histogram of this error is also presented in Figure 3.14. A root mean square error
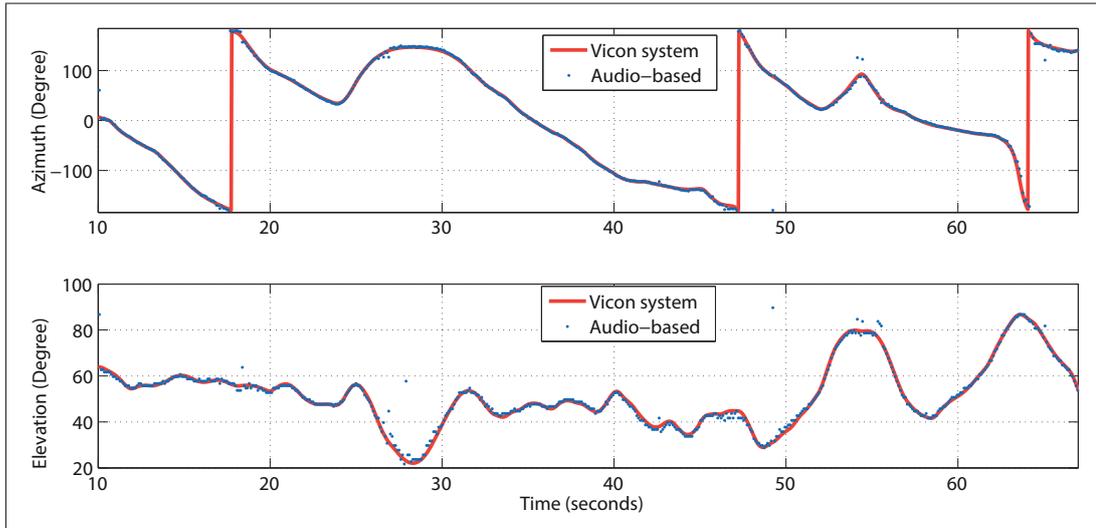
Figure 3.13: The relative azimuth and elevation to a flying pocket size rotor craft MAV, from a stationary MAV on the ground, compared against the true values found using the motion capture system

of 1.39 degrees were computed for the angular error between the true and the audio-based direction estimations.

Experiments were also performed to test the performance of the system and the coherence pruning unit when experiencing the sound from multiple robots. For this, the sound of a flying rotor-craft MAV was recorded in real flight using an onboard digital sound recorder. Different instances of the recorded sound was then played simultaneously from multiple loudspeakers for various positions. The exact position of the loudspeakers and the perceiving robot was measured using the Vicon system. Figure 3.15 illustrates the box plots for the angular error in the cases of two, three and four simultaneously playing robot sounds. In these plots, the source number corresponds to the order of the sound sources that are located from the same time window of acoustic measurements, i.e. The source number n is the source that is localized after $n-1$ pruning iterations. In all the experiments up to three sources were localized with a good accuracy. However, the precision dropped significantly after the third pruning step resulting in poor measurements for the direction of the fourth source. The number of existing sources did not have a major effect on the localization performance for the first three dominant sources, with only a minor increase in the number of outlier.

Note that although the system was only capable of providing reliable direction measurements for the first three dominant targets at a given time instance, however, in reality direction information of more targets can be obtained throughout time. This is because the engine sounds are constantly varying and the robots are moving, which alters the superiority of the targets throughout time, allowing more than three robots to be discovered in time. A time-filter algorithm, such as particle filters, could potentially be employed to keep track of the direction of more robots throughout time.
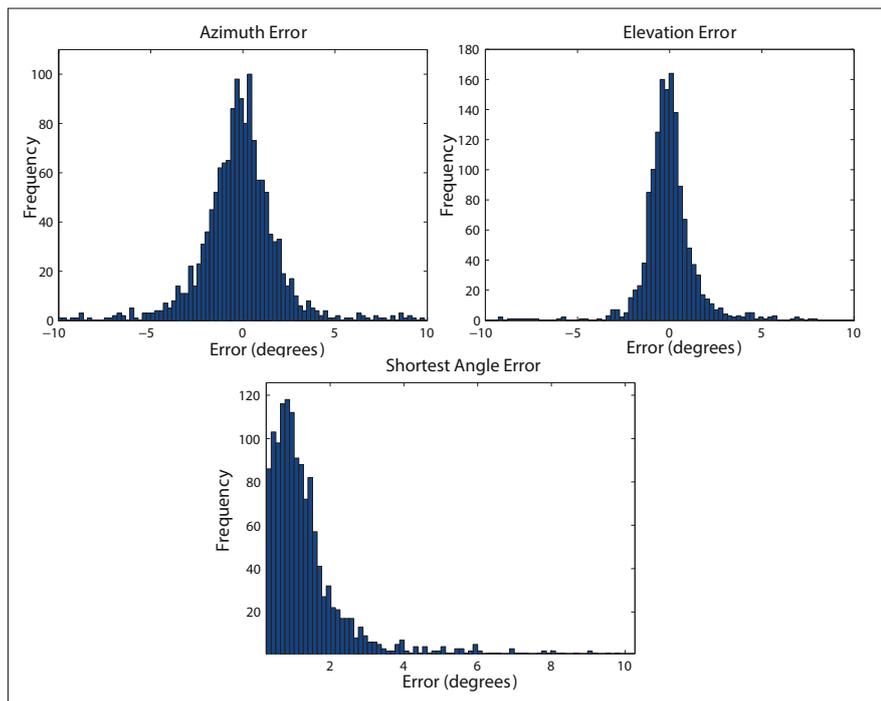
Figure 3.14: Error histograms of the relative azimuth and elevation errors for this experiment

Outdoor experiments were also carried out with the available fixed wing MAVs, where accurate relative direction measurements to a flying MAV, that was manually controlled, was obtained from a stationary MAV on the ground. The used fixed wing MAV had a single small electric motor engine that produced sound pressure levels in the range of 55dB to 72dB depending on the input throttle value. A detection range of beyond 150 meters were observed when the target robot was flying with near full throttle speeds. This range was reduced for lower throttle values. Figure 3.16 shows a plot of audio-based relative azimuth measurements compared against GPS based measurements from one of these experiments illustrating a good correspondence between the two values. The throttle input of the target robot and the inter robot distance is also shown for each measurement.

Furthermore, experiments were executed to try and measure the relative bearing between two flying fixed wing MAVs using the sound of their engines. The robots were programmed to fly autonomously between predefined way-points using their onboard autopilots. The perceiving Robot was programmed to switch its engine off occasionally. However, in these experiments the direction of the target robot was obtained only few times and only when the robots were in a close proximity with each other. Two main reasons were identified for this poor detection range. Firstly, and most evidently, was that these robots required a low input throttle value for their autonomous operations which resulted in very low sound levels generated by the target robot that could hardly be heard even by the human operators. Secondly, the platform vibrations, actuator noises and the high air flow on the microphones, due to the fast speed of these MAVs, led to a lower signal to noise ratio that also influenced the detection range. This
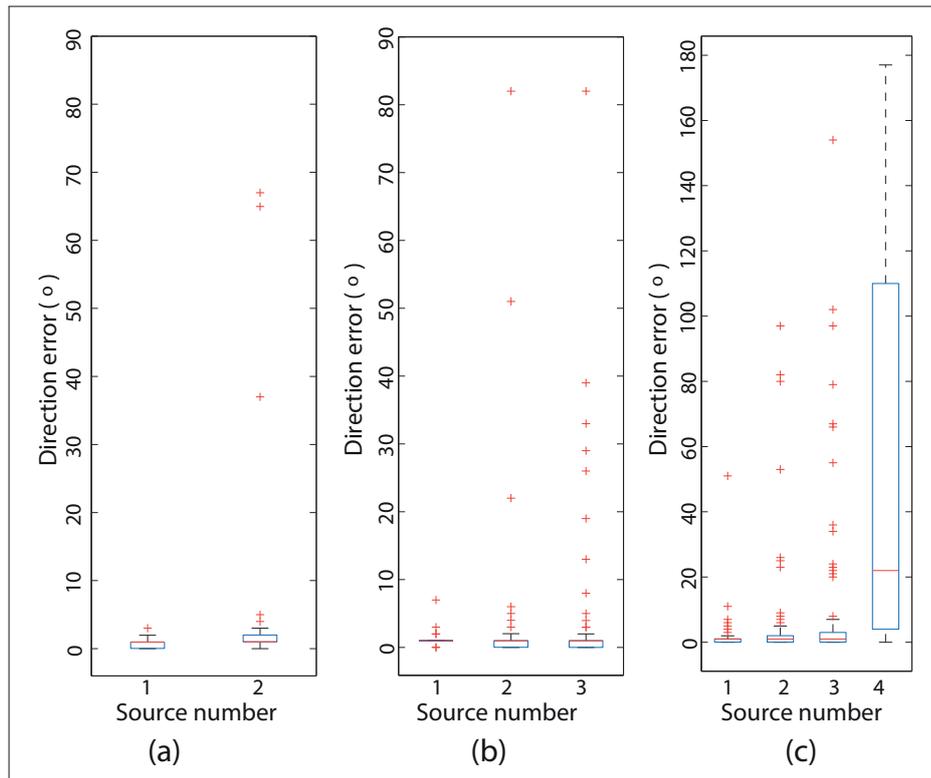
Figure 3.15: Box plots illustrating the statistical properties of the relative direction measurement error in experiments with multiple rotor-craft MAV sounds. a) Two targets b) Three targets c) four targets
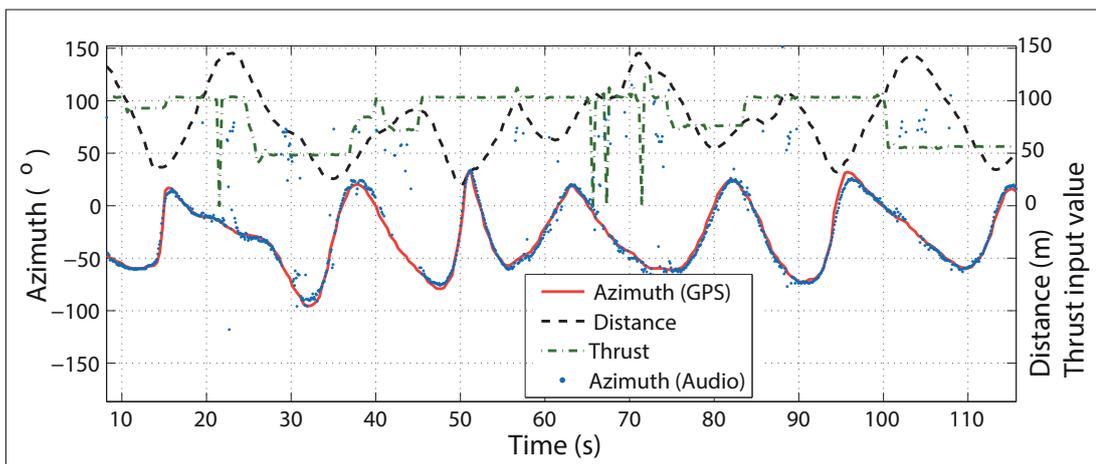
Figure 3.16: The relative azimuth to a flying fixed-wing MAV from a similar but stationary MAV on the ground. The throttle input of the target robot and the inter robot distance computed from GPS positions is also shown for each measurement.

is because these noises are mostly dominant in the lower frequency regions of the sound's spectrum, where the target sound is also mostly present, particularly for longer distances.

3.2.6   Summary of method

In this section an onboard solution for measuring the relative directions between individuals in a team of MAVs, based entirely on the sound generated by their engines, and in the absence of the self-engine noise, was presented. The proposed system exhibited a good accuracy, particularly in low acoustic noise environments. The detection range for the system depends on the sound level of the target engine's and the acoustic noise level, hence making the suitability of the approach dependant on the type of platforms and the operating environments. It was further shown that the direction of up to three local neighbours, demonstrating the highest superiority in the coherence measurements, could be measured at a single time instance by pruning the coherence measurements. Time-filter tracking methods could potentially be employed to track the direction of more robots throughout time as the superiority of robots alter in time. The proposed method has the advantage of energy efficiency due to its passivity that could also provide important bearing information of other non-cooperative aerial robots and platforms that produce sound during flying. The inability to identify the robot identities or to distinguish team-mates from other flying robots, the no self-engine noise constraint, and the dependency of the detection range on the engine speed of the target robots are some of the limitations of this method and motivation for the next section.

3.3   Active audio-based relative positioning system

A passive audio-based relative bearing measurement system, based on perceiving the engine-sound of robots, was described in the previous section. Although this method provides several advantages, some of the mentioned drawbacks motivated research for an alternative solution. This section explains a method based on active acoustic signalling, where individuals generate a unique bird-like sounds, to assist each-other in obtaining the inter-robot relative bearing information. Furthermore, an estimator is derived for the fixed-wing type of MAVs that fuses the bearing measurements with information about the motion of the MAVs, provided by their onboard sensors, to estimate more robustly the relative position, i.e. both relative range and bearing, throughout time. An estimator for rotor-craft type of MAVs could potentially be derived in a similar manner.

Figure 3.17 presents the schematic diagram of the overall audio-based relative positioning system. The system is divided into two main parts of "Target robot" and "Perceiving robot" states to illustrate the main units of the system involved at each state. In the target robot state, the "Chirp Generator" of a robot generates unique chirp sound of predefined rate and frequency. In the perceiving state, sound waves are picked up by an onboard microphone array and are continuously checked by the "Chirp Detection and Separation" unit for existence of chirps in the sound mixture. When a full chirp is detected, it is filtered out from the sound mixture and is
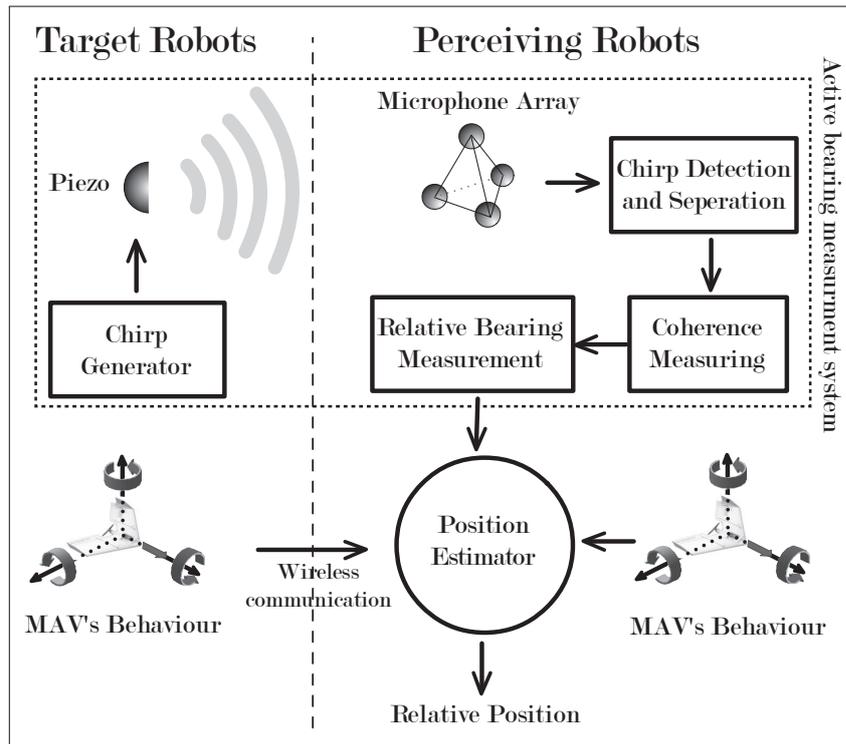
Figure 3.17: Schematic diagram of the proposed relative positioning system illustrating main parts of the system.

then passed to the "Coherence Measuring" unit. The coherence measuring unit, that was previously described in Section 3.2.2, obtains a measure of similarity between the chirps among all of the microphone pairs. The frequency range used in the coherence measuring, defined by f_{min} and f_{max} in Equation (3.6), is chosen as the minimum and maximum frequencies of the perceived chirp. The value of $\alpha = 1$ is used in Equation (3.5) to consider equally all the frequencies within the chirp's range of frequencies. The resulting coherence measurements and the knowledge of the microphone array's geometry is then used by the Relative Bearing Measurement unit, previously explained in Section 3.2.3, to estimate a measure of the target's direction. Finally, the position estimator unit, based on particle filters, estimates robustly the relative location of the target robot by fusing the noisy bearing measurement with information about the relative motion of robots throughout time. The relative motion between robots are computed using information from the onboard proprioceptive sensors and with the aid of a communication network. The particle filter is preferred over a parametric approach, such as the Extended Kalman Filter, due to the non-linear nature of the relative motion dynamics of the MAVs. A more detailed explanation of the units that were not explained in the previous section is presented in the following subsections.
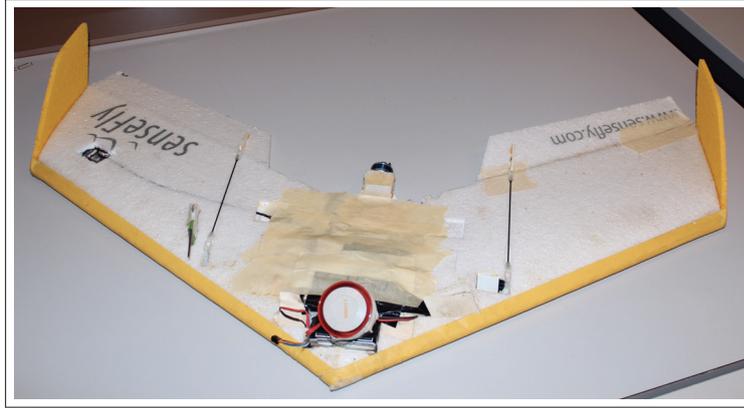
Figure 3.18: A piezo transducer attached near the nose of a fixed wing MAV that is used as the target robot in the experiments described later in Section 3.3.4

3.3.1 Chirp generator

Piezoelectric transducers are simple, inexpensive and lightweight devices that are suitable to be used on MAVs. These devices generate sound by converting electrical signals into mechanical vibrations. A loud sound wave can be produced if the input signal frequencies are close to the resonance frequency of the piezo element. In addition, focusing the available electrical power into a narrow range of frequencies results in a longer range transmission of the sound waves. Hence, to generate a loud sound wave that is needed for achieving a longer detection range, particularly suitable for outdoor operations, and to avoid the problem of ambiguous bearing measurements related with localization of narrowband sounds, described in Section 2.2, a piezo transducer is used on the robots to generate band limited chirp signals. Figure 3.18 shows one of the fixed wing robots equipped with a piezo transducer that is used in the experiments described later in Section 3.3.4.

The chirp generating unit of every target robot generates linear chirps with a predefined and unique chirp rate. Since an entire chirp is used by the perceiving robots for computing a single bearing measurement, the time interval between the chirps can be chosen in accordance with the required measurement rates. To generate continuous chirps, the frequency of the input sine-wave signal to the piezoelectric element is varied in time:

$$f(t) = F_{str} + \frac{F_{end} - F_{str}}{\Delta T} \bmod \left(\frac{t}{\Delta T} \right) \tag{3.13}$$

where F_{str} and F_{end} are the starting and ending frequencies of the chirp that is chosen differently for every robot, and ΔT is the chirp duration. Figure 3.19 illustrates the sound wave and spectrogram of an in-flight sound recording, performed with three flying fixed-wing MAVs, i.e. one perceiving MAV and two target MAVs emitting periodical chirps. .
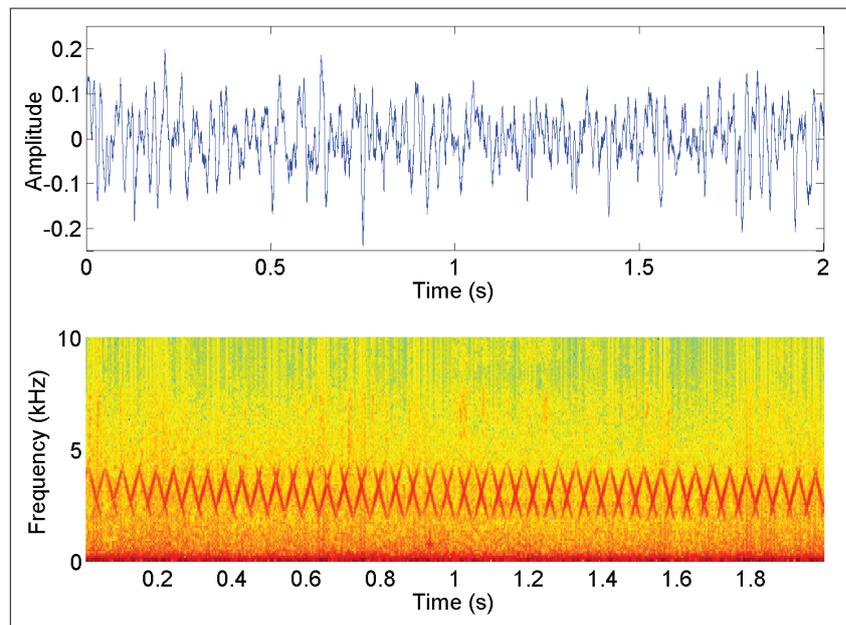
Figure 3.19: Sound wave and spectrogram of an in-flight sound recording involving one perceiving robot and two chirping MAVs. The two linear chirps are in the same frequency band and have a different chirp rate. One of the target robots produced chirps with up-sweep frequency from 1700kHz to 4700kHz and the other target robot produced linear chirps with down-sweep frequency from 4700kHz to 1700kHz.

3.3.2 Chirp detection and extraction

This unit is responsible for the detection and extraction of a chirp in the perceived sound mixture. For this purpose, one of the microphone signals are continuously checked by the unit to find the existence of chirps in the sound mixture. The presence of a chirp, belonging to a specific target robot, is detected by template matching technique, where a continuous cross correlation of the sound mixture with a template of the desired chirp determines the existence and the time segment containing the entire chirp.

After a chirp is detected, it is filtered out from other sounds and overlapping chirps that might also be present in the selected time segment. For this purpose, initially the time segment from all of the microphone signals are passed through a band-pass filter to remove the unwanted low and high frequency noises that are outside the chirp's frequency range (see Figure 3.21.b). Furthermore, Fractional Fourier transform (FRFT) [Namias, 1980] is used to remove the noises within the frequency range of the chirp.

First proposed by Namias [Namias, 1980], FRFT has been recently favoured in the field of signal processing [Ozaktas et al., 1994], mostly for recovering signals from noise [Erden et al., 1999, Kutay et al., 1997] and particularly when dealing with chirp signals [Durak and Aldirmaz, 2010]. Unlike Fast Fourier transform (FFT), the FRFT provides a compact representation of chirp signals, which makes it possible to remove the noise inside the same frequency region as
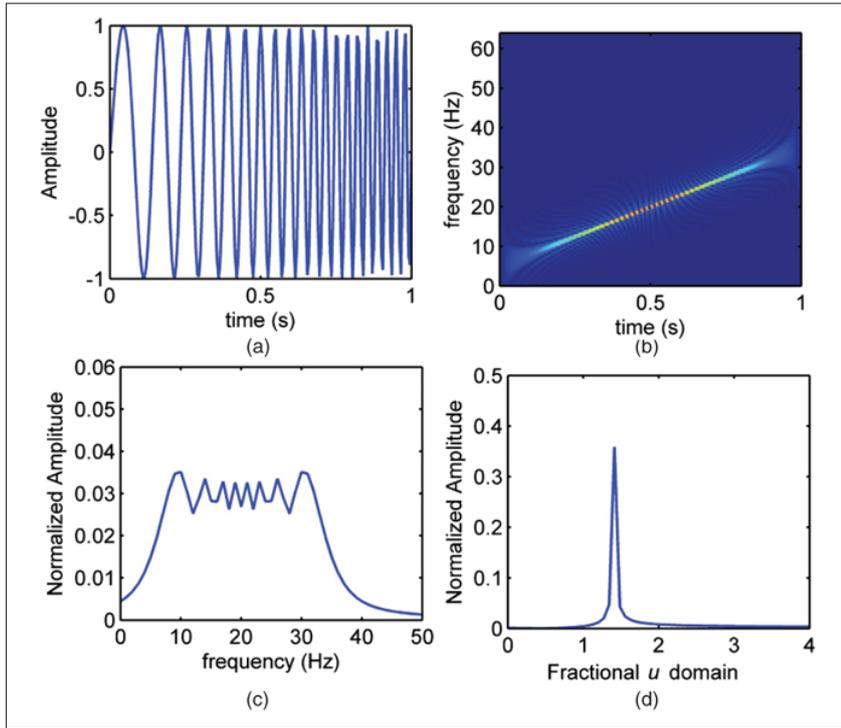
Figure 3.20: A comparison between the time, frequency and fractional domain representation of a linear chirp. a) Time domain b) Spectrogram c) Frequency domain d) Fractional domain

the chirp, that cannot be removed with traditional frequency domain filters. FRFT, transforms the linear chirp to an intermediate domain between the time and frequency to have the chirp represented by a single sharp peak. Figure 3.20 illustrate a comparison between the time, frequency and fractional domain of a linear chirp. In simple words, FRFT projects the signal, represented in time-frequency plane, onto a line of arbitrary angle φ defined by the FRFT order α, whereas the FFT projects the signal on to the y-axis, i.e. $\varphi = \pi/2$. A method for computing the fractional Fourier transform by means of fast Fourier transform algorithm is presented by [Garcia et al., 1996].

To represent the detected chirp in its most compact form, the Fractional Fourier transform (FRFT) of the time window containing the entire chirp is computed with an FRFT order of α obtained by the following equation.

$$\alpha = \frac{2}{\pi}\varphi = \frac{2}{\pi}\tan^{-1}\left(\frac{f_s}{F_{end} - F_{str}}\right) \tag{3.14}$$

where F_{str} and F_{end} are the starting and ending frequencies of the perceived chirp in Hertz and f_s is the sampling frequency. Equation (3.14) was derived from the geometrical relationship between the chirp rate and FRFT order for time-frequency discretized chirp signals provided by [Capus and Brown, 2003].
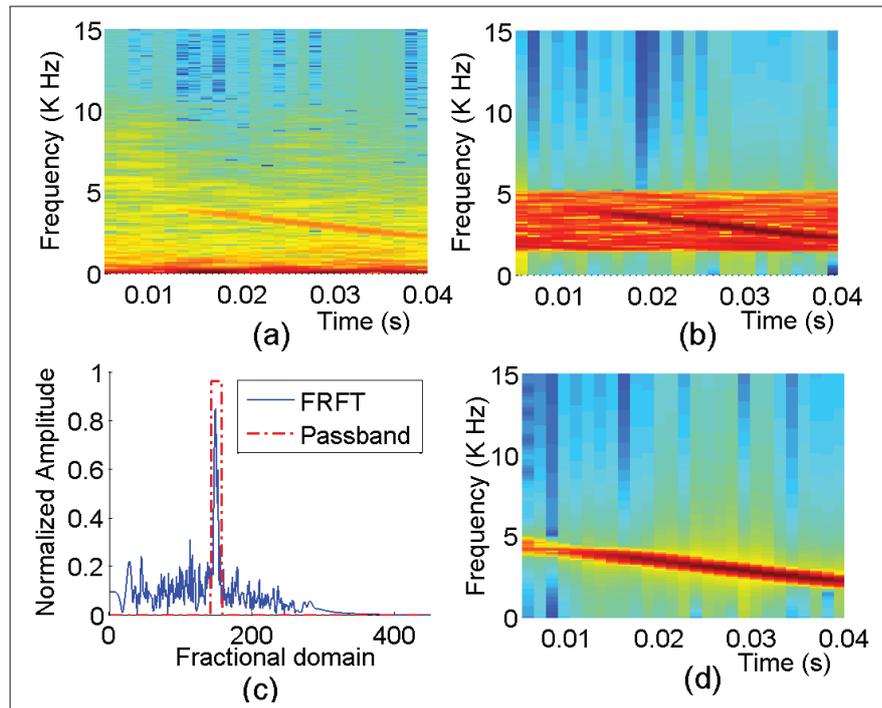
Figure 3.21: In-flight sound of a chirping MAV recorded by an observing MAV in different steps of the chirp extraction procedure (a) Spectrogram of a detected chirp (b)Spectrogram of the signal after band-pass filtering (c) FRFT transform of the band-passed chirp and the corresponding passband region (d) Spectrogram of the final filtered chirp.

Upon computing the FRFT, the result contains an impulse-shaped peak that corresponds to the desired chirp. The chirp is then filtered out from other sounds, by only retaining the bin with the highest peak along with its few nearby bins and setting all other bins to zero (illustrated in Figure 3.21.c) The ratio of the peak value to the mean value of all zeroed bins prior to zeroing provides a good measure for the quality of the perceived chirp, i.e. signal to noise ratio. This measure is computed and used later as a reliability measure for the obtained bearing measurements, where only measurements satisfying a predefined reliability level are used in the update step of the particle filter. Furthermore, the filtered chirp in the FRFT domain is transformed back to the time domain by computing the inverse FRFT. Steps and result of the chirp extraction procedure for a detected chirp in a real world experiment with two flying MAVs is illustrated in Figure 3.21.

3.3.3 Position estimator

The already described units of Figure 3.17, provide a method for allowing individuals to obtain an instantaneous relative bearing measurement to sound emitting target robots in the neighbourhood, that might be noisy or unavailable at times. It is now required to estimate more reliably the relative bearing information and to also obtain some information about the

relative range of the target robots. This is achieved by fusing the relative bearing measurements, that are obtained throughout time, with the relative motion dynamics of the perceiving and the target MAVs, measured by their onboard proprioceptive sensors. The particle filtering method, that was initially introduced in Section 2.3.2, is used here for fusing this information in order to recursively estimate the probability density of the target location.

At time instant k, the relative position of a single target robot is modelled using a set of N particles of vectors $p_i(k)$ and weight $w_i(k)$, where $p_i(k) = (p_{xi}(k), p_{yi}(k), p_{zi}(k))$ is a vector in the body-fixed coordinate system of the perceiving robot that starts at its origin and ends at a point in space. $p_i(k)$ can also be described in the body-fixed spherical coordinate system $(\angle\phi, \angle\theta, r)$ by:

$$u_i(k) = (\phi_i(k), \theta_i(k), r_i(k)) \qquad i = 1, 2, ..N \tag{3.15}$$

where ϕ_i is the relative azimuth defined in the range $[-\pi, \pi]$, θ_i is the relative elevation defined in the range $[-\pi/2, \pi/2]$ and r_i is the relative range defined in the range $[R_{min}, R_{max}]$. R_{min} and R_{max} are dependent on the platform size, and the sound pressure level generated by the piezo transducer, respectively. For the MAVs and the piezos used in this work these ranges are found approximately to be $[1, 250]$ meters.

A three dimensional state vector is specified for every particle:

$$S_i(k) = \begin{bmatrix} \phi_i(k) & \theta_i(k) & r_i(k) \end{bmatrix} \tag{3.16}$$

The unit starts by forming an initial set of particles $\{S_i(0), i = 1 : N\}$ for every target robot detected to be in the neighbourhood. Particles either could be distributed uniformly over the entire state space, or only over a desired part of the state space if some prior knowledge about the possible location of the target is available. In this work, to reduce the number of required particles, the initial state space is reduced to all vectors in the space having a small deviation from the first reliable bearing measurement.

As described previously in Section 2.3.2, a particle filter estimator consists of two main steps that is repeated iteratively: Prediction and Update.

3.3.3.1 Prediction step

In the prediction step, a set of new particles $\{\tilde{S}_i(k)\}$ is predicted by propagating $S_i(k-1)$ according to a probabilistic relative motion model. This model is derived with the assumption that, at every time step, robots have a forward motion, (i.e. along the x axis of their body-fixed coordinate system), followed by a three dimensional rotation, (i.e. yaw (λ), pitch (β) and roll (α) rotations around the z, y and x axis of the body fixed coordinate system respectively). Figure 3.22 illustrates the positions of two robots in two successive time steps consisting of a perceiving robot A and a target robot B. Using linear algebra the following relationship
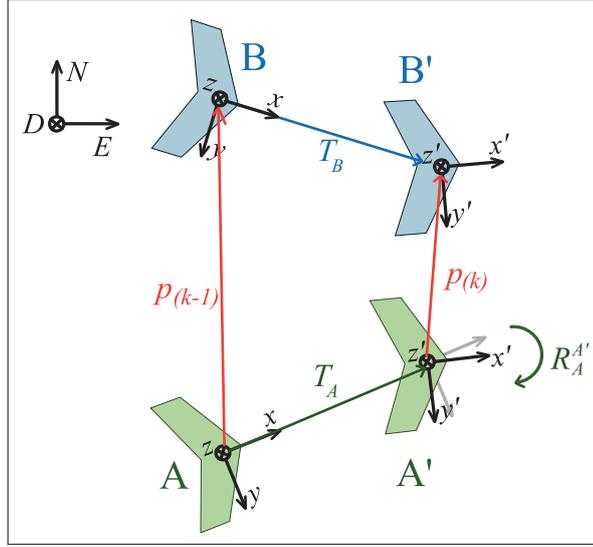
Figure 3.22: Positions of two robots (perceiving A and target B) in two successive time steps along with coordinate systems and connecting vectors

between the vectors can be described:

$$\overrightarrow{p}(k) = R_{B'}^{A'}(R_B^{B'}\overrightarrow{T}_B) + R_A^{A'}\overrightarrow{p}(k-1) - R_A^{A'}\overrightarrow{T}_A \tag{3.17}$$

where R_I^J is a rotation matrix that rotates a vector from the coordinate system I to the coordinate system J:

$$R_I^J = R_z(\lambda_J - \lambda_I).R_y(\beta_J - \beta_I).R_x(\alpha_J - \alpha_I) \tag{3.18}$$

$(\lambda_I, \beta_I, \alpha_I)$ is the bearing of the coordinate system I relative to a fixed NED coordinate system and (R_z, R_y, R_x) are basic rotation matrices that rotate vectors about the local z, y and x axis respectively.

Equation (3.17) is used by the perceiving robot A to predict the vector $\overrightarrow{p}_i(k)$ for particle i from its previous value $\overrightarrow{p}_i(k-1)$. For this, speed and orientation measurements of the perceiving robot, and of the target robot that are transmitted via a communication network, are used. The forward motion vectors $\overrightarrow{T}_A$ and $\overrightarrow{T}_B$ are initially computed from the speed sensor readings V_A and V_B measured at time instance $k-1$

$$\overrightarrow{T}_A = \begin{bmatrix} (V_A(k-1) + \xi_V)\, dt \\ 0 \\ 0 \end{bmatrix} \qquad \overrightarrow{T}_B = \begin{bmatrix} (V_B(k-1) + \xi_V)\, dt \\ 0 \\ 0 \end{bmatrix} \tag{3.19}$$

where dt is the time interval between the two time steps and $\xi_V = N(0, \sigma_V)$ is a random number generated with a normal distribution of mean zero and standard deviation σ_V. The value of σ_V is chosen in relation with the reliability in the speed sensor reading measurements.

Furthermore, the rotation matrices $R_I^J(\lambda + \xi_\lambda, \beta + \xi_\beta, \alpha + \xi_\alpha)$ in Equation (3.17) are computed from the bearing measurements $(\lambda, \beta, \alpha)_{I,J}$ and using Equation (3.18), with $\xi_\lambda = N(0, \sigma_\lambda)$, $\xi_\beta = N(0, \sigma_\beta)$ and $\xi_\alpha = N(0, \sigma_\alpha)$ to simulate the noise in the measurements. Finally, Equation (3.17) can be solved for the prediction $\vec{p}_i(k)$ of particle i.

3.3.3.2   Update step

In the update step, the likelihood of the obtained audio-based relative bearing measurement is investigated for every particle and particles are weighted according to this measure. For this investigation, we propose the likelihood function:

$$w_i = \frac{1}{\sigma_m \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\varepsilon_i}{\sigma_m}\right)^2} \tag{3.20}$$

where

$$\varepsilon_i = \angle(\vec{b}_m(k), \vec{p}_i(k)) \tag{3.21}$$

is the angle between the measured bearing $\vec{b}_m(k)$ at time k and the predicted vector $\vec{p}_i(k)$ of particle i. The value of σ_m reflects the confidence of the bearing measurements and is found empirically. As mentioned in Section 3.3.2, only reliable measurements obtained from chirps with a good signal to noise ratio, i.e. that satisfy the predefined reliability level, is used in the update step. The update step is skipped for unreliable measurements. In addition to this, a simple outlier rejection method further investigates whether the update step should be performed, where newly arrived measurements with large deviation from the predicted particle distribution are rejected.

Note that, the likelihood function (3.20) is formed by assuming that the angular error between the direction measurements $\vec{b}_m$ and the true directions $\vec{b}_T$, have a Gaussian distribution with mean zero and standard deviation σ_m. i.e.

$$\angle(\vec{b}_m, \vec{b}_T) \sim \mathcal{N}(0, \sigma_m^2) \tag{3.22}$$

Experiments showed that this assumption approximately holds. Figure 3.23 shows the relative bearing measurements and the histogram of the angular errors from a field experiment involving a target robot and a perceiving robot.

3.3.3.3   Relative Position Estimation

The relative range and bearing of the target can be estimated at each time step from the probability density function represented by a particle set. For this, a weighted mean of all particles' positions could be employed. However, to avoid inaccurate estimations for situations
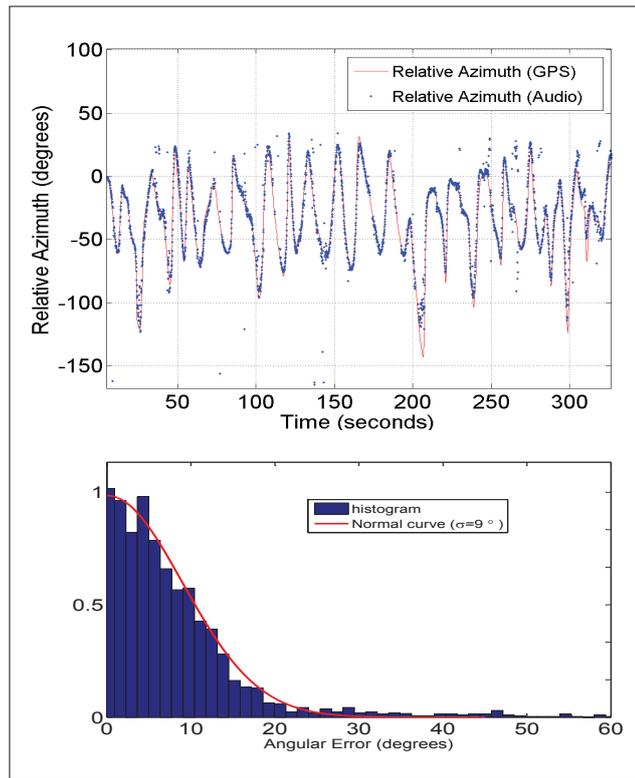
Figure 3.23: Audio-based bearing measurements from an experiment involving a perceiving robot and a target robot. In this experiment, the perceiving robot was fixed over the ground to eliminate the uncertainties, in the estimation of the true bearings, caused by the onboard gyroscopes and the GPS of the perceiving robot. The target robot is flown manually in proximity of the perceiving robot and its onboard GPS is used only to compute the true bearings. top: Audio-based and GPS based relative azimuth measurements. bottom: histogram of the angular errors between the measured bearings and the true bearings.

with multi-modal distributions, a weighted mean of particles located in a local neighbourhood of the particle with the highest weight is used instead:

$$\bar{S}_T = \sum_{i=1}^{K} w_i S_i : \forall |S_i - S_{max}| < \xi \tag{3.23}$$

Finally, the particles are resampled according to their normalized weights to avoid the problem of degeneracy of the particle filtering algorithm.
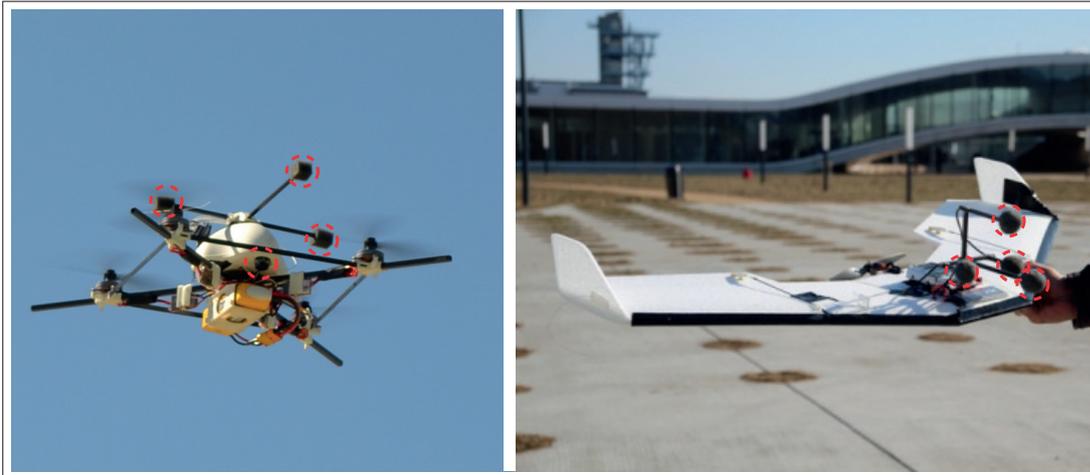
Figure 3.24: Picture of the two type of MAV platforms used in the experiments, indicating the position of the four onboard microphones. Both robots were designed and developed in the Laboratory of Intelligent Systems at EPFL. More information about the fixed wing robot can be found in the work of [Leven et al., 2007]

3.3.4   Experiments and results

3.3.4.1   Relative bearing measurement

Initially, experiments were undertaken to test the proposed active audio-based relative bearing measurement system. For this, real world experiments with both fixed-wing and rotorcraft type of MAVs were performed. In these experiments, the target robots were equipped with a piezo transducer and a small electrical circuit based on a micro-controller for generating periodic linear chirps. According to the datasheet of the used piezo element, a maximum sound output of 100dB at 1 meter is produced at the resonance frequency. On the other hand, the perceiving robots were equipped with an array of four microphones as shown in Figure 3.24. The microphones were covered with a small piece of foam for wind protection and for reducing the noise introduced by the airflow caused by the motion of the robots.

Figure 3.25 illustrate results of an outdoor experiment where a perceiving rotorcraft MAV was flown manually in the vicinity of a continuously chirping rotorcraft MAV. The target MAV generated continuous linear chirps, each chirp with a down-sweep frequency from 4700kHz to 1700kHz in approximately 25.6 milliseconds. The microphone array on the perceiving rotorcraft MAV composed of three microphones forming a triangle of edge length 18cm, each placed between the propellers, and the fourth microphone was placed approximately 5.2cm under the center point of the triangle. A sampling rate of 40kHz was used for simultaneously sampling the four microphone signals by the onboard microcontroller. Throughout the experiment, the audio-based relative azimuth and elevation measurements from the audio processing board, the robot's attitude measurements provided by the onboard inertial measurement unit (IMU), and the GPS positions, were transmitted and stored on to a ground station through an XBee

communication link. The stored GPS and IMU measurements were used later to compute the GPS-based relative bearing measurements for comparison. Figure 3.25(a,b) shows plots of audio-based and GPS-based relative azimuth and elevation measurements for parts of this experiment illustrating a good correspondence between the two values. A histogram of the angular difference between the GPS based and audio-based relative bearing measurements for this experiment is shown in Figure 3.25.c. Experiments with the used rotor-craft MAV showed the success of detecting, filtering and localizing the chirp sound of target robots in presence of the self-engine noise produced by the four engines of the robot. A detection range of ≈ 60 meters were found in these experiments.
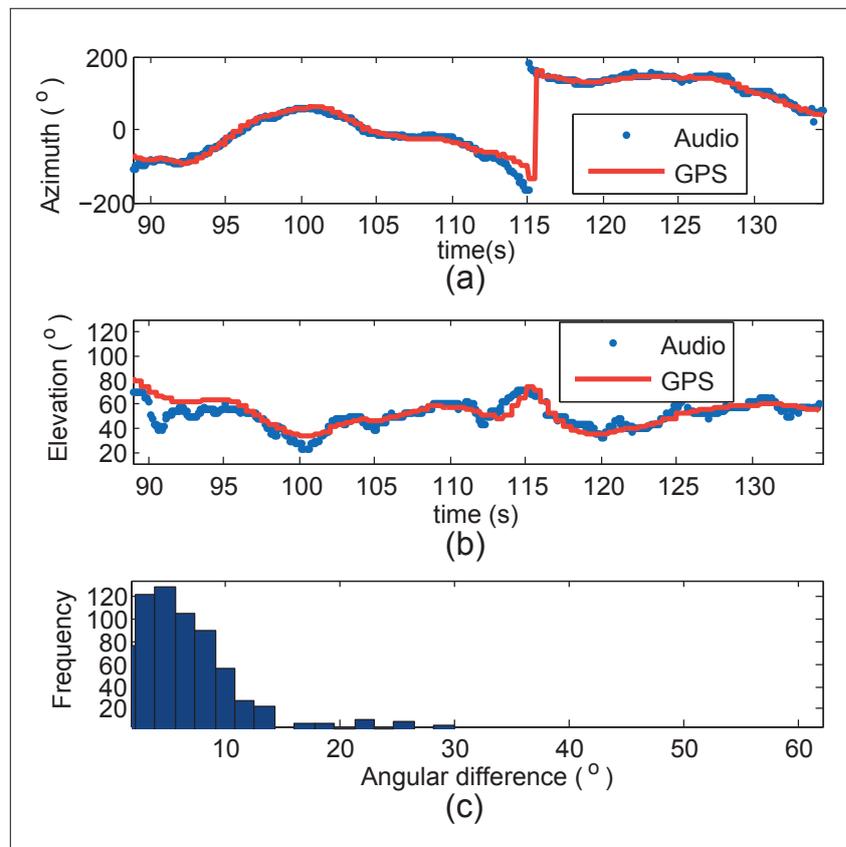
Figure 3.25: Relative bearing measurements from a flying rotorcraft MAV to a chirping team-mate.

In addition, a distributed leader-follower motion coordination between the two rotorcraft MAVs were implemented to further test the onboard relative bearing measurement system. For this purpose, a reactive controller was implemented on one of the robots to autonomously follow a chirping leader robot based entirely on the the instantaneous 3D audio-based bearing measurements. The heading of the follower robot was directly controlled by the relative azimuth measurements to face the direction of the leader robot, the relative elevation measurements were used to control the speed of the robot to reach the target, and the values

Figure 3.26: Photo showing an instance of the audio-based leader-follower experiments involving two rotorcraft MAVs, where a fully autonomous follower MAV used the on-board active relative bearing measurement system to reactively follow a teleoperated chirping leader robot. Videos can be found at http://lis.epfl.ch/ABSMAV
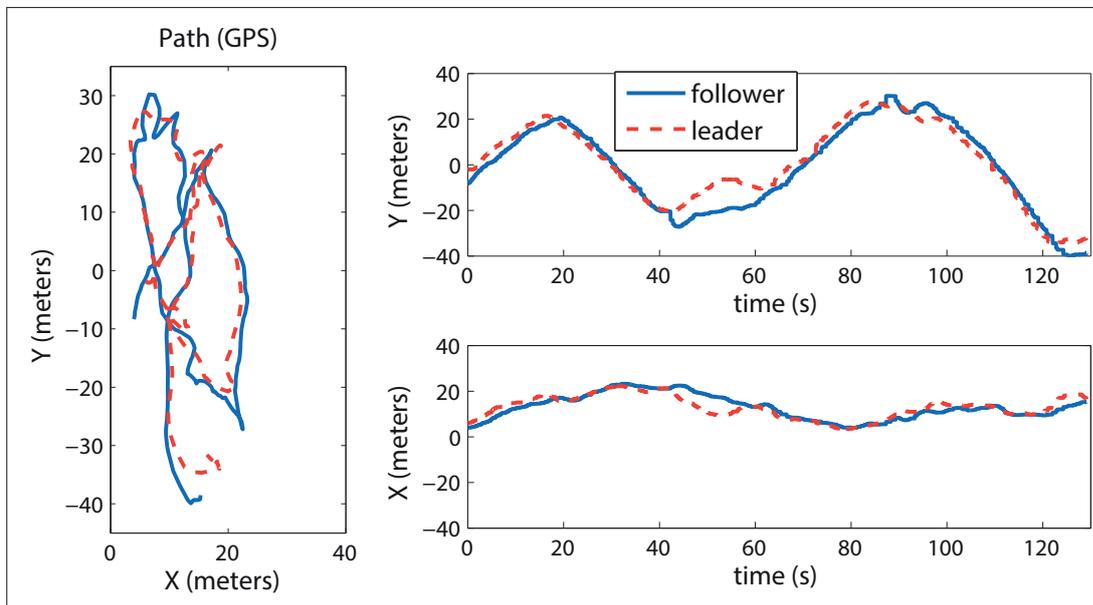
Figure 3.27: Path and positions of the MAVs, extracted from the GPS logs of the leader-follower experiment, illustrating the success of the fully autonomous follower robot to reactively follow the chirping leader robot using entirely the locally perceived sound waves

from an onboard pressure sensor were used to maintain the altitude of the robot. The overall controller can be thought of as a potential vector field in space, with vectors guiding the robot to maintain its current altitude while pointing towards the goal point of |elevation| = 90 degrees (i.e. the point exactly above the leader robot if the leader is at a lower altitude, or exactly bellow the leader robot if it is at a higher altitude). Figure 3.26 shows an instance of the experiment where the autonomous follower MAV is following a teleoperated chirping leader robot. Figure 3.27 shows the path and positions of the MAVs, extracted from the GPS logs of one of the experiments, illustrating the success of the fully autonomous follower robot to follow the leader robot while entirely depending on the 3D bearing information obtained from the locally perceived sound waves.

Furthermore, real world experiments involving three fixed-wing flying robots were carried out. In these experiment a perceiving robot was flown manually in the vicinity of two simultaneously chirping robots. The target robots were controlled automatically by their onboard autopilots based on a set of predefined GPS way-points. One of the target robots produced linear chirps with an up-sweep frequency from 1700kHz to 4700kHz and the other robot produced linear chirps with down-sweep frequency from 4700kHz to 1700kHz. The rate of chirping for both robots were set to about 20 chirps per second with each chirp having a duration of approximately 0.05 seconds. An instance of the sound recording by the perceiving robot is shown in Figure 3.19 that illustrates the chirps from the two targets robot in the sound mixture. The microphone array on the perceiving robot formed a regular tetrahedron of edge length 10 cm and a sampling rate of 48kHz was used for this experiment. The array's position and dimension was selected in order to prevent the drag caused by the microphones from affecting the MAV's stability. The orientation, altitude, air-speed and global positioning information of the MAVs were measured using onboard sensors and transmitted and recorded on the ground station through a wireless communication network. The roll and pitch orientations of the MAVs were measured using onboard gyroscopes, and since no compass sensor was present on these MAVs, the heading information from the onboard GPS sensor was used instead. The MAVs were controlled to fly within the visual range of a safety pilot while the engine power of the perceiving robot was occasionally reduced to improve the detection range by increasing the signal to noise ratio. Plots of audio-based and GPS based relative azimuth measurements along with the GPS based relative distance is shown in Figure 3.28. Simultaneous bearing measurements to the two robots is obtained indicating the success of the system in correctly identifying and separating the two overlapping chirps. Experiments also indicated that the maximum detection range of the system depends on the self-motor sound of the perceiving robot, where a lower detection range was obtained for high motor speeds. This can be seen in Figure 3.28.b, at around $t = 5$ seconds where a higher motor input results in less reliable measurements for the further away robot and only the closer robot is continuously localized. This is because, as the engine-speed is increased, the noise at higher frequencies where the chirps are expected is also increased that results in the reduction of the signal to noise ratio. A detection range of up to ≈ 200 meters were obtained by the perceiving robots when gliding with the engine turned off.
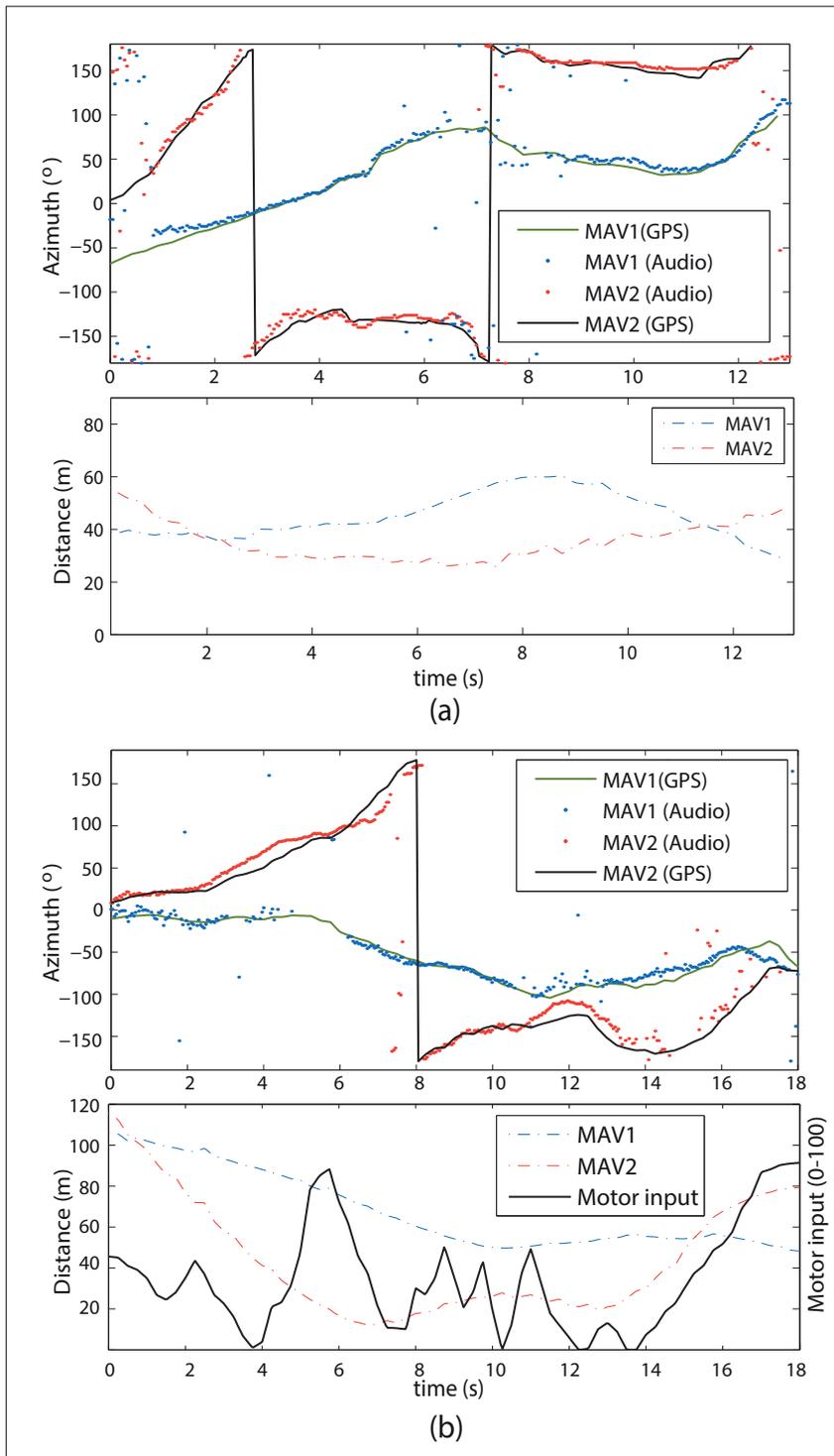
Figure 3.28: Results of a real world experiment involving three flying fixed-wing robots, where the relative direction of two simultaneously chirping robots is measured from a perceiving robot. (a) and (b) illustrate two different instances of the experiment, each containing plots of relative audio-based azimuth measurements and GPS-based azimuth estimates. The relative distance between the robots shown is computed using GPS positions.

In addition, experiments were performed in simulation to test the performance of the method in cases with more than two overlapping chirps. Note that in practice it is possible to reduce the number of overlapping chirps, by reducing the chirp rate of the robots, or by using different piezos and spreading the chirps in different frequency ranges. In these experiments, the sound wave at each microphone position were simulated by superimposing the waves from randomly positioned point sources, based on the principle of superposition of sound waves. Simple free-field sound-wave propagation model was used to simulate the sound wave of a source at the point of the microphone. Each source generated a continuous chirp sequence, with unique chirp rates but with equal chirp duration, to ensure that all the chirps overlap with each other. The chirp sequence for every source was randomly started in every experiment to simulate the fact that the chirp timing between robots are not synchronous and that chirps from different robots could overlap at different times. A sampling frequency of 40kHz, chirp duration of 50ms and a microphone array with the same dimension as the one on the rotorcraft MAV was used for the simulations. Figure 3.29 shows the spectrogram of a microphone signal and the relative bearing measurements from an experiment with 10 overlapping chirps. It can be seen that all chirps are successfully extracted from the sound mixture and the directions of all 10 targets are identified correctly. Experiments showed that for large number of overlapping chirps, and since the chirp timing between the robots are not synchronous, multiple overlapping times could occur on a single chirp that might prevent it from being detected or correctly localized. Figure 3.30 shows the number of sources that was incorrectly localized for the cases of 6,8,10,12 and 14 overlapping chirps, each computed from 100 experiment runs with random overlapping configurations.

3.3.4.2 Relative position estimation

The experimental data obtained from the in-flight experiments involving three fixed-wing MAVs were employed to test the performance of the particle filter estimator in estimating both the relative bearing and the relative range of the two target robots. It was observed that the estimator was capable of tracking the relative bearing of robots, and also provide an estimate of the relative range, with a good accuracy. Figure 3.31 shows the path of all three robots, recorded by the GPS sensors, for 25 seconds duration of flight time. The relative bearing measurements and estimates for this duration of time is shown in Figure 3.32. It can be observed that, the relative bearing of the targets are tracked correctly even at times that there are no reliable observations available. The relative range estimations along with the particle distributions and GPS based range estimates are shown in Figure 3.33. It can be seen that, the particles gradually converge towards the correct relative range and furthermore track it with an acceptable accuracy. As expected, the speed of convergence and the accuracy in the relative range estimations are dependent on the motions and positions of the robots. For some types of relative motions, the particles having an inaccurate range are eliminated faster than for other types of motions. Figure 3.33 shows that in the first few seconds, where the perceiving robot is further away from the target robots, and robots are moving towards each other, particles are still widely spread in relative range although they have converged to the
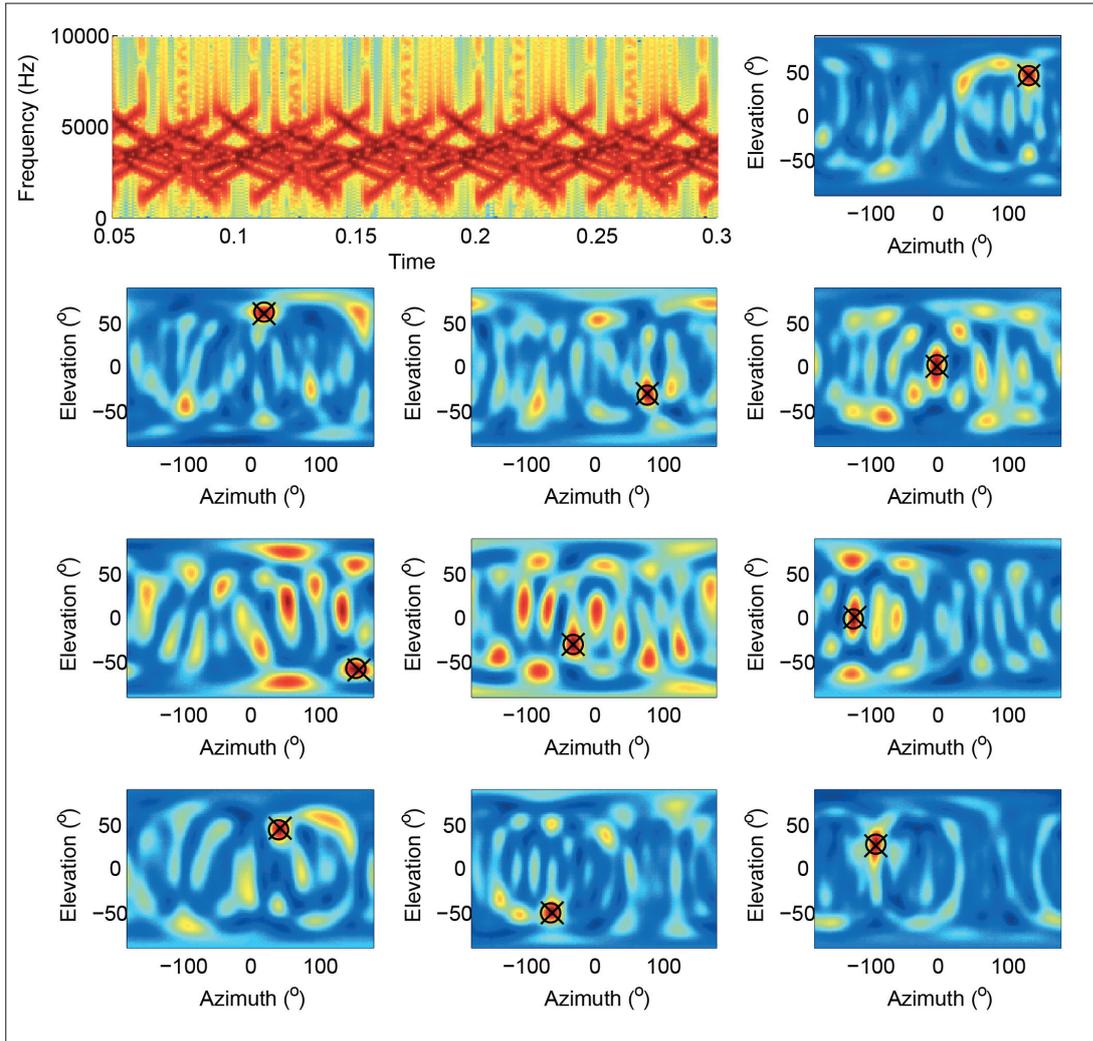
Figure 3.29: The spectrogram of a microphone signal and the measured direction likelihoods of the extracted chirp signals in a simulation experiment with 10 simultaneously chirping targets. The true source direction is marked with a cross and the direction with the maximum likelihood is marked with a circle.
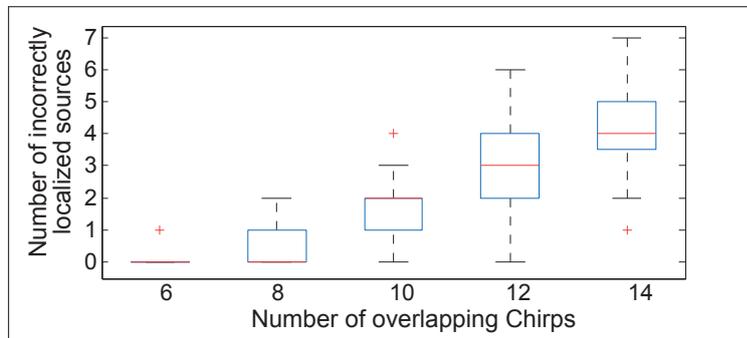
Figure 3.30: Number of incorrectly localized sources in a single time frame for the cases of 6,8,10, 12 and 14 overlapping chirps, computed from 100 random overlapping configurations.
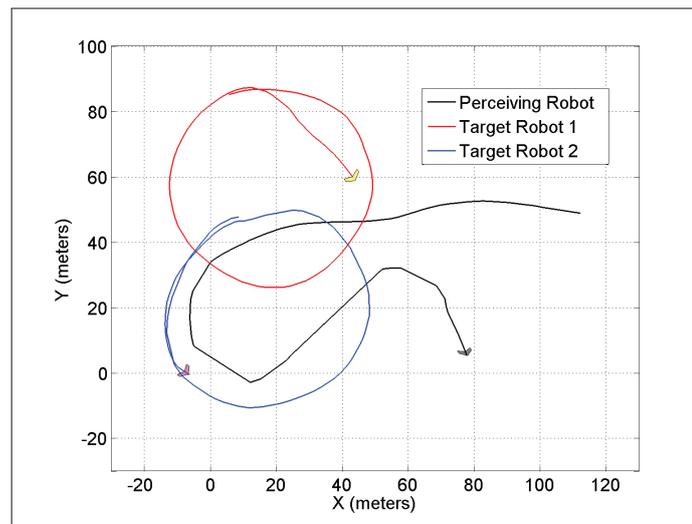
Figure 3.31: The motion path of robots recorded by onboard GPS sensors, for 25 seconds of flight time in an experiment involving one perceiving robot to locate and track two target robots

correct bearing. As the robots get closer and pass each other, the disparity of the particles in range is also reduced.

3.4 Conclusion

In this chapter, two methods for measuring the relative bearing of individuals in a group of MAVs were presented. Both methods were based on a small, compact and onboard microphone array for passively measuring the sound waves in space and estimating the direction of arrival of sound waves. The first method relied on the sound waves generated by the engine of other robots to obtain the relative bearing measurements. This method exhibited a good direction accuracy in the absence of the self engine noise, but the detection range was strongly dependant on the output sound level of the targets and the acoustic noise level of the environment. Furthermore, by pruning the coherence measurements it was possible to simultaneously measure the direction of up to three target robots having the highest sound superiority. The method had the advantage of energy efficiency due to its complete passivity and the capability of measuring the direction of other non-cooperative sound emitting aerial platforms. However, the inability to identify the robot identities, the no self-engine noise constraint and the dependency of the detection range on the engine speed of the target robots were some of the limitations of this method motivated the need for the second approach. The second method was based on perceiving acoustic chirp signals emitted by the robots using an onboard piezo transducer. The unique and high frequency chirps allowed the robots to obtain the relative bearing measurements in presence of the self-engine noise and to identify the identity of robots for each measurement. Although the chirping rate and frequency could be selected to reduce or remove overlapping chirps in the sound mixture, however experiments
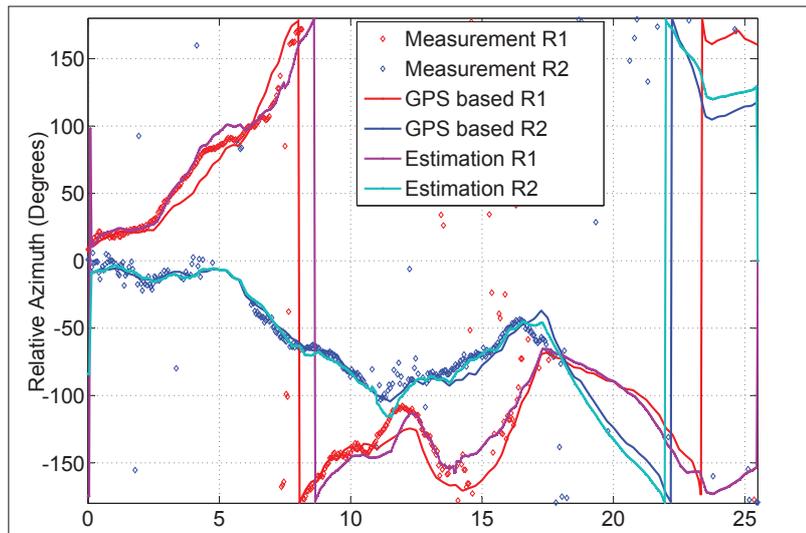
Figure 3.32: The relative azimuth estimates from the Position Estimator unit compared with the relative azimuth computed from the GPS positions and the onboard IMU orientations. The measurements obtained by the Relative Bearing Measurement unit are also shown by small markers.
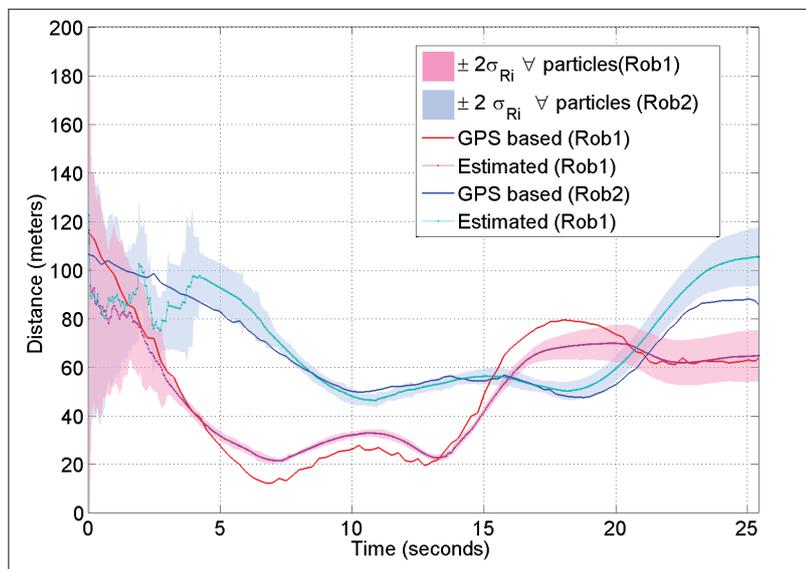
Figure 3.33: The Relative range estimations, standard deviation of the relative range of all particles and GPS based range values

showed that with the described fractional Fourier filter it is possible to extract and localize many overlapping chirps. Furthermore, an estimator was derived for the fixed-wing type of robots to allow individuals to estimate and track the relative position of other members throughout time, by fusing the bearing measurements with the state information of the MAVs provided by the onboard sensors. In the next section, a solution to self localization of individuals, all relative to a single reference point, is described based on the proposed relative bearing measurement system described in this chapter.

4 Audio-based localization for swarms of micro air vehicles

L OCALIZATION is one of the key challenges that needs to be considered beforehand to design truly autonomous MAV teams. Individual's knowledge about their three dimensional position is required for allowing MAVS to autonomously navigate to different points in space and for performing many aerial coverage tasks such as exploration and mapping. Furthermore, sharing this information with other robots through a communication network provides an alternative solution to the inter-robot relative positioning problem that was discussed in previous chapter.

In this chapter, we present a cooperative method to address the localization problem for a team of MAVs, that is independent of any external systems while satisfying the constraint of MAVs, where individuals obtain their position through perceiving a sound-emitting beacon MAV which is flying relative to a reference point in the environment. In particular, we provide a solution for a team of fixed wing robots that accommodates the motion constraints of these type of robots. The method is based on the on-board audio-based relative bearing measurement system, described in the previous chapter, for allowing robots to measure the relative direction to the sound emitting beacon robot. This information, along with the internal sensory information and prior knowledge about the behaviour of the beacon robot, is used by the individuals to localize themselves and the beacon robot without the need for a communication network. The proposed method is evaluated both in simulation and in real world experiments.

4.1   Introduction

MAV self localization problem, is the problem of estimating the three dimensional position of an MAV relative to a single reference point by the MAV itself . In Section 1.1, the importance of self-localization for designing truly autonomous MAV teams were described. It was described that, individual's knowledge about their 3D location is necessary for allowing robots to autonomously navigate to different points in space and to achieve many aerial coverage tasks such as exploration and mapping. This information could further be used to avoid inter robot collisions by priori spatial separation of individuals at different altitudes or locations [Allred et al., 2007, Cole et al., 2008]. In addition, by sharing their positions with other team members, individuals can obtain the relative positioning information, allowing them to form formations [Basiri et al., 2010] [Moshtagh et al., 2009] and avoid collisions with other MAVs without the need of priori spatial separation [Carnie et al., 2006].

The constraints of MAVs limit the transfer of successful localization solutions from ground robots and large aerial vehicles to the small scale MAVs. Section 1.2.1 provided an overview of the existing localization methods used for MAVs, where solutions were divided into two main categories of Global methods and Local methods. Global localization methods rely on external systems, such as the 3-D motion tracking cameras or global positioning system (GPS) satellites, that might not always be accessible to the MAVs. On the other hand, local localization methods, such as the vision based SLAM, had the important advantage of independency, however, they required high computational power and data storage, to detect, localize and keep track of features in the environment, that might not be available on the small scale and inexpensive MAVS. The need for real-time processing of high resolution and high frame-rate images, the dependency on illumination, visual contrast, weather conditions and the limited field of view of vision sensors, the errors caused due to the high or insufficient number of features in the images, the long displacement between loop closings and the fast dynamic nature of MAVs, were some of the other drawbacks of the visual SLAM methods for aerial robots [Artieda et al., 2009].

The aim of the work presented in this chapter was to obtain a localization solution for a team of MAVs that is independent of any external system while avoiding the high complexity and drawbacks of the local localization methods. For this purpose, we propose a cooperative method allowing robots to obtain their 3D positions through cooperation, where a robot in the group acts as a positioning beacon assisting others with obtaining their positions. The method is based on the on-board audio-based relative bearing measurement system that was described in Section 3.3, that allows robots to measure the relative bearing to a sound emitting beacon robot. The robots then use this information, along with their on-board sensory information to localize themselves all relative to a single reference point. Particularly we derive a solution for fixed wing type of MAVs that can accommodate the motion constraint of these type of robots and their requirement to always maintain a forward speed for staying airborne.

This chapter is organized as follows: Section 4.2 describes our proposed method for self localization of MAVs in a Multi-MAV system, where an estimator is derived for estimating the robot locations throughout time based on the bearing-only measurements to the beacon robot. Section 4.3 provides the results of simulation and real world experiments of the proposed method. Section 4.3 describes a conclusion to this chapter and potential future works.

4.2 Proposed method

This section explains our method for localizing MAVs in a Multi-MAV system. This method can be considered as a combination between the two class of localization approaches, described in Section 1.2.1, to employ some of the advantages from both classes. It allows a group of MAVs to cooperatively localize themselves using only their onboard sensors, and independent of any external systems, while avoiding the high complexity nature of the local localization methods. The idea here is that a single MAV in the group starts to fly in a circular pattern, acting as a positioning beacon and attracting the attention of other robots by emitting sound waves, that are continuous sequence of acoustic chirps of predefined rate and frequency. All MAVs are equipped with the audio-based relative bearing measurement system presented in Section 3.3 to passively perceive the chirps and measure the direction of the beacon robot. Upon hearing the calls of the beacon robot, MAVs start to estimate their positions and the position of the beacon robot simultaneously, all relative to a single reference point, throughout time. An Extended Kalman Filter (EKF) estimator is derived for this estimation that is explained in detail in Section 4.2.1 . No communication between robots is required as robots only consider the prior knowledge of the beacon robot's behaviour in their estimations.

As illustrated in Figure 4.1, the beacon robot is controlled to circle around a desired reference point, while trying to maintain a previously defined altitude, speed and circling radius. Many control strategies for guiding an MAV on a circular path exists [Frew et al., 2007, Nelson et al., 2007]. In this work, a vector field based controller, similar to the one proposed in [Nelson et al., 2007], was used to control the motion of the beacon MAV around the reference point. The beacon MAV can consider a static point on the ground that is detected by an onboard camera [Grocholsky et al., 2006], or a static acoustic target on the ground that is detected by an onboard microphone array, similar to our proposed solution in Chapter 2, as the reference point.

4.2.1 Position estimation using bearing-only measurements

A position estimator, based on the Extended Kalman Filtering (EKF) method [Welch and Bishop, 1995], is derived to allow observing robots to robustly estimate their locations throughout time. This is achieved by fusing the noisy relative bearing measurements, with information about the motion of the robot itself, gathered by the onboard proprioceptive sensors, and taking into account the prior knowledge about the behaviour of the beacon robot. The estimator is recursive and consists of an Initialization step and two iterative steps, Prediction and Update,
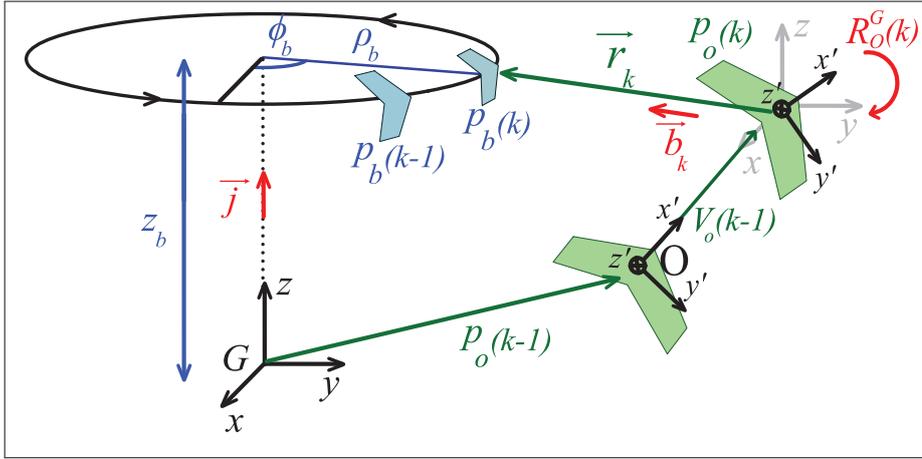
Figure 4.1: Diagram illustrating the positions of two MAVs, beacon MAV (p_b) and observing MAV (p_o), for two successive time steps, and introducing some other symbols that are used for describing the position estimation method.

that are explained in the following subsections.

At time instant k, the position of the beacon MAV and an observing MAV, relative to the reference point, is given by position vectors $p_b(k)$ and $p_o(k)$ respectively, where p_b is defined in Cylindrical coordinate system by $p_b = (\rho_b, \phi_b, z_b)$ and p_o is defined in Cartesian coordinate system by $p_o = (x_o, y_o, z_o)$. The combination of both position vectors is considered as the state vector X for the EKF:

$$X = \left(\begin{array}{c} p_b \\ p_o \end{array} \right) = \left[\begin{array}{cccccc} \rho_b & \phi_b & z_b & x_o & y_o & z_o \end{array} \right]^T \tag{4.1}$$

Furthermore, a 6×6 covariance matrix $P(k)$ defines the state error covariance matrix at time instant k.

Initialization

In this work, an initial state estimation strategy is proposed in order to obtain a good initial guess of the state vector $X(0)$ to have a faster convergence in the state estimation. The EKF is initialized after the first reliable bearing measurement is obtained, by using the MAV's orientation and altitude sensor values:

$$
\begin{aligned}
p_b(0) &= (R_b, 0, Z_b) \tag{4.2} \\
p_o(0) &= Z_b \vec{j} - \ell(R_O^G \vec{b_0}) \tag{4.3} \\
\ell &= \begin{cases} \dfrac{(Z_b - Z_o)}{sign(Z_b - Z_o(0)) \vec{j} \cdot R_O^G(0) \vec{b_0}} & Z_b \neq Z_o \\ \dfrac{D_M}{2} & Z_b = Z_o \end{cases}
\end{aligned}
$$

where, Z_b and R_b are the prior knowledge of the beacon MAV's altitude and circling radius respectively, $\vec{j}$ is a unit vector along the positive z axis of the global coordinate system G, $(\cdot)$ is the vector dot product, $Z_o(0)$ is the measured altitude of the observer MAV at time zero and D_M is the maximum detection range of the bearing measuring sensor. Vector $\vec{b_0}$ is a unit vector in the observer MAV's body fixed coordinate system O pointing along the direction of the initial relative bearing measurement. $R_O^G(0)$ is a rotation matrix that rotates vectors from coordinate system O to G:

$$R_O^G(k) = R_z(-\lambda_o(k))R_y(-\beta_o(k))R_x(-\alpha_o(k)) \tag{4.4}$$

$(\lambda_o(k), \beta_o(k), \alpha_o(k))$ are the yaw, roll and pitch orientation measurements of the observer robot and (R_z, R_y, R_x) are basic rotation matrices that rotate vectors about the local z, y, x axis respectively. In the case of ideal measurements and when $Z_b \neq Z_o$, Equation (4.3), obtained using basic vector operations, calculates the center point of a circle of radius R_b that have the observer MAV on its circumference. Furthermore, a covariance matrix $P(0)$ is initialized:

$$P(0) = diag(\sigma^2_{\rho_b(0)}, \sigma^2_{\phi_b(0)}, \sigma^2_{z_b(0)}, \sigma^2_{x_o(0)}, \sigma^2_{y_o(0)}, \sigma^2_{z_o(0)})$$

where $\sigma^2_{x(0)}$ is the initial covariance of the state variable x that are chosen in accordance to the reliability of sensor readings and the uncertainties in the initial state estimation.

Prediction

In the prediction step, the current state of the system $\tilde{X}(k)$ is predicted from $X(k-1)$. For the observer MAV, a probabilistic motion model and the onboard sensor information, providing the speed and orientation of the MAV, is used to predict the position vector $\tilde{p}_o(k)$ from $p_o(k-1)$.

$$\tilde{p}_o(k) = p_o(k-1) + R_O^G(k-1) \begin{bmatrix} V_o(k-1)dt \\ 0 \\ 0 \end{bmatrix} \tag{4.5}$$

where $V_o(k)$ is the speed sensor reading and dt is the time interval between the two time steps. The motion model (4.5) is derived by assuming that, at every iteration, the MAV has a forward motion along the x-axis of its body fixed coordinate system, followed by a three dimensional rotation.

If communication between the robots were available, the speed and orientation values of the beacon MAV along with conversions between Cylindrical and Cartesian coordinate systems could also be used to predict the beacon MAV's position vector $\tilde{p}_b(k)$. However, as we are interested in a solution that does not depend on a communication network, only the prior knowledge about the speed V_b and circling radius R_b is used to obtain $\tilde{p}_b(k)$:

$$\tilde{p}_b(k) = p_b(k-1) + \begin{bmatrix} 0 & \frac{V_b}{R_b}dt & 0 \end{bmatrix}^T \tag{4.6}$$

Furthermore, a prediction of the state covariance matrix, $\tilde{P}(k)$, is obtained by assuming that the uncertainty in state predictors (4.5) and (4.6) is a zero mean multivariate Gaussian.

Update

In the Update step, the relative bearing measurement, presented by vector $\vec{b}_k$, is used to update the state prediction $\tilde{X}(k)$. For this, a measurement model to predict the relative bearing from the state predictions is defined:

$$\theta = \tan^{-1}\left(\frac{r_y}{r_x}\right) \qquad \varphi = \tan^{-1}\left(\frac{r_z}{\sqrt{r_x{}^2 + r_y{}^2}}\right) \tag{4.7}$$

where

$$\vec{r} = (r_x, r_y, r_z) = \left(\rho_b \cos\phi_b - x_o, \rho_b \sin\phi_b - y_o, z_b - z_o\right)$$

is a vector in the global coordinate system G that starts at the position p_o and ends at the position p_b. θ and φ are the azimuth and elevation of vector $\vec{r}$. The predicted bearing $(\tilde{\theta}, \tilde{\varphi})$ is found by substituting the state predictions $\tilde{X}(k)$, from Equations (4.5) and (4.6), into Equation (4.7).

Furthermore, an innovation $\mu(k)$ is defined as the difference between the predicted bearing $(\tilde{\theta}_k, \tilde{\varphi}_k)$ and the measured bearing $(\hat{\theta}_k, \hat{\varphi}_k)$:

$$\mu(k) = \begin{bmatrix} \hat{\theta}_k - \tilde{\theta}_k & \hat{\varphi}_k - \tilde{\varphi}_k \end{bmatrix}^T \tag{4.8}$$

where $\hat{\theta}_k$ and $\hat{\varphi}_k$ are the azimuth and elevation of vector $\vec{b}_k$ expressed in the coordinate system G, i.e $R_O^G(k)\vec{b}_k$. The innovation covariance matrix $S(k)$ is computed by:

$$S(k) = H\tilde{P}(k)H^T + D \tag{4.9}$$

where D is the error covariance of bearing measurements and is found empirically. H is the Jacobian of the measurement model (4.7) with respect to the states:

$$H = \begin{bmatrix} \frac{\partial \theta}{\partial X} \\ \frac{\partial \varphi}{\partial X} \end{bmatrix}\Bigg|_{\tilde{X}(k)} = \begin{bmatrix} H_{11} & \dots & H_{16} \\ H_{21} & \dots & H_{26} \end{bmatrix}\Bigg|_{\tilde{X}(k)} \tag{4.10}$$

where

$$H_{11} = \left(y_o \cos\varphi_b - x_o \sin\varphi_b\right)/\Re_1$$
$$H_{12} = \rho_b \left(\rho_b - x_o \cos\varphi_b - y_o \sin\varphi_b\right)/\Re_1$$
$$H_{14} = \left(-y_o + \rho_b \sin\varphi_b\right)/\Re_1$$
$$H_{15} = \left(x_o + \rho_b \cos\varphi_b\right)/\Re_1$$
$$H_{13} = H_{16} = 0$$
$$H_{21} = \left((z_o - z_b)\left(\rho_b - x_o \cos\varphi_b - y_o \sin\varphi_b\right)\right)/\Re_3$$
$$H_{22} = -\rho_b (z_o - z_b)\left(y_o \cos\varphi_b - x_o \sin\varphi_b\right)/\Re_3$$
$$H_{24} = (z_o - z_b)\left(x_o - \rho_b \cos\varphi_b\right)/\Re_3$$
$$H_{25} = (z_o - z_b)\left(y_o - \rho_b \sin\varphi_b\right)/\Re_3$$
$$H_{23} = -H_{26} = \sqrt{\Re_1}\Big/\Re_2$$
$$\Re_1 = {\rho_b}^2 + {x_o}^2 + {y_o}^2 - 2\rho_b\left(x_o \cos\varphi_b + y_o \sin\varphi_b\right)$$
$$\Re_2 = \Re_1 + (z_o - z_b)^2$$
$$\Re_3 = \Re_2 \sqrt{\left(x_o - \rho_b \cos\varphi_b\right)^2 + \left(y_o - \rho_b \sin\varphi_b\right)^2}$$

Finally the states are updated:

$$X(k) = \tilde{X}(k) + K(k)\mu(k)$$
$$P(k) = \tilde{P}(k) - K(k)H\tilde{P}(k)$$

where $K(k)$ is the Kalman gain at time k derived by:

$$K(k) = \tilde{P}(k)H^T S(k)^{-1}$$

4.3 Experiments and results

To verify the proposed bearing-only position estimator, initially experiments were performed in simulation using a group of modelled MAVs. Simulated MAVs were presented by a first order 3D flight model with three degrees of freedom for the airspeed, turn rate and the altitude, all controlled by PID controllers. The MAV's airspeed, turn rate and altitude dynamics had rate limitations and were influenced by a uniform noise. Furthermore, the sensors that provide the MAV's orientations, speed, altitude and the relative bearing to the beacon MAV were modelled to be affected by a zero mean uniform noise while the relative bearing sensor was also limited in range. Model parameters were tuned to best represent the simple MAV platform that was used throughout the real experiments [Leven et al., 2007]. A vector field controller was used on the MAVs to steer their motions onto a circular path around desired waypoints. For the beacon robot this waypoint was always a fixed point in space, while for the observer robots random waypoints were generated sequentially to navigate them between random points in the space. Figure 4.2 shows the results of multiple simulation runs involving a beacon MAV and an observer MAV. It shows the gradual convergence of the position estimations to the true position and the reduction in the error covariance, once the beacon MAV is within the detection range of the observer MAV. The absolute estimation error for all the EKF states in
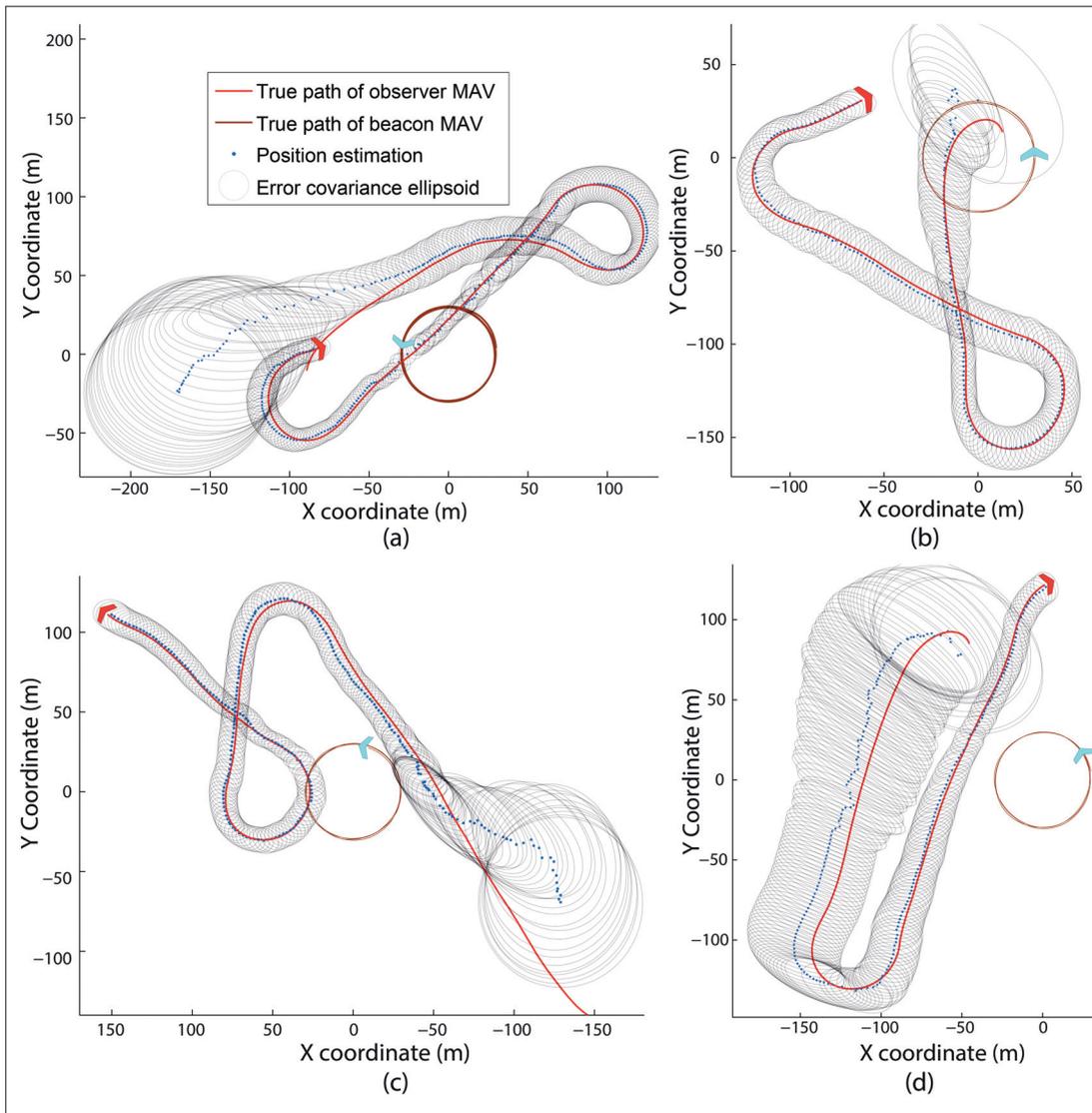
Figure 4.2: Results of position estimations in multiple simulation runs involving 1 observer and 1 beacon MAV. A uniform noise of $\pm 5^o$ for relative-bearing/attitude sensors and $\pm 1m$ for altitude/speed sensors is used in the simulations. The bearing sensor's detection range is defined as $150m$. Plots compare the position estimates alongside the true path of the robots showing the convergence of the estimates to the true position.

one of the experiments is shown in Figure 4.3. The gradual reduction of errors in the position estimates corresponding to the beacon robot, and the observer robot itself, is illustrated in this figure. As expected, the convergence speed depends on the relative motion between the robots, where for the motions that result in a faster change of the relative bearing a faster localization is obtained. Upon localization, a good position tracking is achieved by the estimator.

Multiple real experiments were performed to further test the proposed method and the localization performance using the available fixed-wing platforms. A beacon MAV, equipped
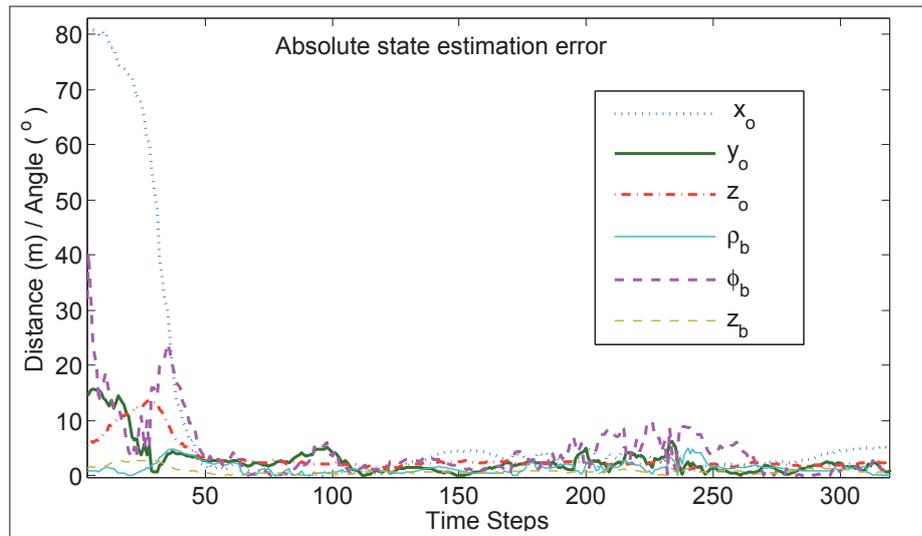
Figure 4.3: Absolute estimation errors of all the EKF states in the experiment of Figure 4.2.a with an observer MAV, indicated by the subscript (*o*), and a beacon MAV, indicated by the subscript (*b*); illustrating the gradual reduction of the position errors towards zero.

with an autopilot, was programmed to fly in circles around a GPS coordinate with constant velocity and constant altitude while emitting chirps using an on-board piezo transducer. An observing MAV, as shown in Figure 3.24, equipped with the on-board audio-based relative bearing measurement system described in Section 3.3, was then flown manually in proximity of the beacon MAV to measure the direction of the incoming acoustic chirp signals. The engine power of the observer MAV was occasionally reduced or turned off to increase the chirp to noise ratio and the detection range. The orientation, altitude, air-speed and global positioning information of both MAVs were measured using on-board sensors and were transmitted and stored on to a ground station. Figure 4.4 shows the audio-based position estimates of the observer robot, along with the path of the robot provided by the on-board GPS sensors, for two instances of the experiments. A good coherence between the GPS-based and audio-based estimates is observed. Figure 4.5 shows the absolute difference between the GPS-based and audio-based state estimates illustrating the convergence of position estimates towards the GPS positions.

4.4 Conclusion

A solution to the problem of MAV swarm localization was presented. This solution consists of a single beacon MAV that circles around a reference point in space while emitting continuous linear chirps of predefined frequency spectrum to assist other MAVs in localizing themselves. MAVs were equipped with the on-board audio-based relative positioning system described in Section 3.3, to measure the bearing to a chirping beacon MAV, and on-board sensors, to obtain information about their motions throughout time. The proposed EKF-based filter was shown
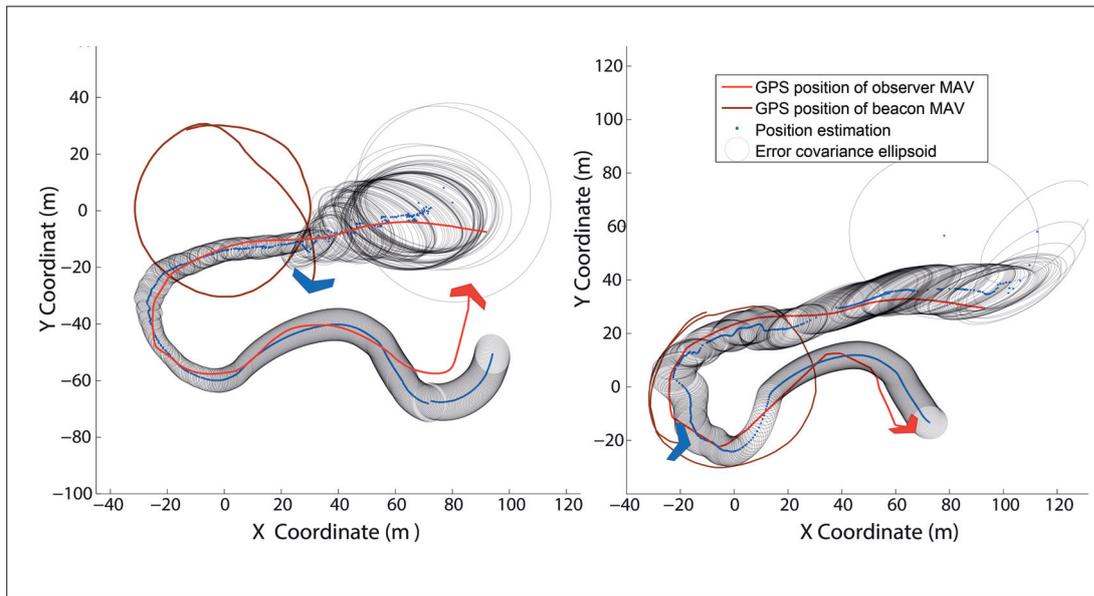
Figure 4.4: Result of a real experiment with two MAVs showing the audio-based position estimates, the error covariance ellipsoids and the path of the MAVs provided by the GPS sensors
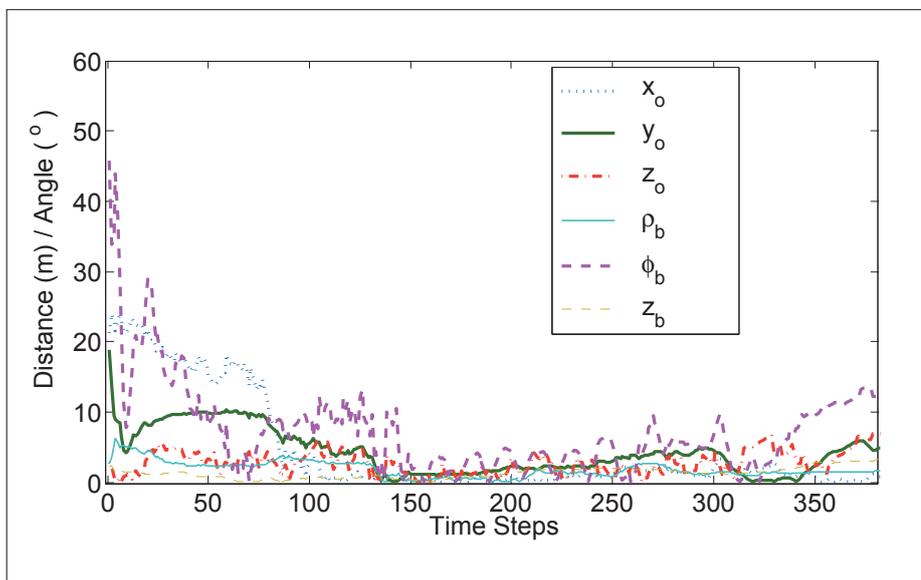
Figure 4.5: Absolute estimation error of all the EKF states in a real world experiment for an observer MAV, indicated by subscript (o), and a beacon MAV, indicated by subscript (b), illustrating the gradual reduction of the state errors, corresponding to the position errors of both robot.

to be well-suited for the sensor fusion and achieving a robust localization. No communication between the robots was required for this purpose and only prior knowledge about robot's behaviours were used in the estimations. Investigating different types of MAV motions that could result in a faster localization, employing multiple beacon MAVs to improve localization performance, and study of switching protocols to switch MAVs between beacon and observer states, for exploration and reduction in the swarm's overall localization error, are some of the potential areas for future explorations.

5 Conclusion

5.1 Main accomplishments

THE Goal of this thesis was to contribute to the field of aerial robotics by proposing potential solutions to some of the challenges faced in the design of autonomous teams of micro aerial vehicles (MAVs), and assisting with their future deployment for real missions. This thesis focused on the new paradigm of exploiting sound waves to obtain independent and on-board solutions that are compatible with the strict constraints of these small-scale, light-weight and inexpensive robots. We proposed novel methods based on an on-board audio-based sensor suite in order to allow individuals inside an MAV swarm to obtain awareness about the position of themselves, the position of other swarm members and to detect and locate other important acoustic targets from the air.

In this thesis, the problem of relative positioning between individuals in an MAV team was considered. Although relative positioning is considered by both the natural and artificial swarm researchers as the essential and sufficient requirement to many swarming behaviours, however, the lack of technological possibilities that could provide individuals with this information, while satisfying the strict constraint of MAVs, have limited solutions to be dependent on external systems that are either impractical or not always available. In this thesis we proposed the naturally inspired solution of using sound waves to obtain an extremely lightweight, small, real-time and on-board system for inter-robot relative positioning of MAV teams. For this purpose, we initially proposed a method based on few spatially separated microphone sensors that allowed robots to measure the relative bearing of their neighbouring robots by perceiving the sound emitted from their engines. The method was shown to provide accurate measurements in the absence of the self engine noise, however, the detection range was dependant on the target robot's engine sound levels and the noise of the operating environment. We then proposed an alternative method where individuals generated unique chirp signals to assist others in obtaining their bearing measurements, which allowed distinguishing the robot identities and the operation in the presence of self engine noises. A filtering method based on fractional Fourier transform allowed the robots to extract a chirp from other overlapping

chirps or noise in the sound mixture and for obtaining accurate localization. An estimator was then derived for the fixed-wing type of robots to allow them to robustly estimate both the relative range and the relative bearings of other robots throughout time, by fusing the audio-based bearing measurements with other on-board sensory information.

Furthermore, a solution to the self-localization problem of individuals in an MAV swarm was introduced, that is independent of any external systems, does not require inter-robot communication, and it can accommodate motion constraints of fixed wing flying robots. The solution, based on the principle of cooperation, consisted of a sound emitting beacon MAV flying in a circular pattern and observing MAVs equipped with the proposed audio-based relative bearing measurement system. An EKF based estimator was shown to be well suited for fusing the audio-based bearing measurements and other on-board sensory information, to obtain the three dimensional positions throughout time.

In addition, this thesis showed how the on-board audio-based sensor suite could also be employed to obtain information about important acoustic targets in the environment. The novel idea of localizing emergency sound sources from airborne micro aerial vehicles was presented that could be a crucial asset particularly in search and rescue operations. Employing a team of MAVs for locating distress sound signals, such as the sound of a person blowing into a safety whistle, could allow the fast localization of victims and the coordination of rescue members in night time and through fog, dust, smoke and foliage. We furthermore proposed multiple solutions to overcome the ambiguity associated with localizing emergency sound signals, or other narrowband sounds in general, and to accommodate different type of aerial robotic platforms. Exploiting the Doppler shift in the sound frequencies due to the motion of the robot, active control of the robot's behaviour while fusing acoustics and attitude information, and modulating the frequency of the sound source itself, were the basis for the three different proposed methods. Experiments with real flying robots showed the success of detecting and correctly localizing different type of emergency sources and verified the proposed methods and solutions.

5.2 Potential applications

This thesis presented multiple practical methodologies designed to enable autonomous operation of groups of small scale aerial robots for real mission scenarios. The ability of obtaining inter-robot relative positioning and self localization information, independent of any external systems, provides a practical solution for realisation of multi-MAV control algorithms, and opens the door to many potential aerial coverage applications. Such abilities would particularly be desired in applications that require fast deployment of groups of aerial robots for operating in unprepared environments. Rapidly deployable communication network, aerial surveillance system, search and rescue operations and environmental monitoring are some of the examples envisioned for truly autonomous MAV swarms.

Furthermore, this thesis presented the idea and methodology of providing micro aerial vehicles

with the important sense of hearing, for detection and localization of different type of acoustic sources in the environment. Apart from the discussed advantages of using sound for achieving spatial coordination in multi MAV systems, audio-based flying robots offer additional avenues to research and applications:

Sense and avoid system: An important requirement for allowing MAVs to operate freely in the airspace is the ability to autonomously sense and avoid mid-air collision with other aircraft and non-cooperative aerial robots. Since MAVs do not have the necessary power and payload to employ radar and other active anti-collision systems, an audio based sense and avoid system could be a promising solution to passively detect and avoid collision with many aerial platforms through the sound emitted by their engines. Many available examples of hear and avoid behaviour in nature [Miller and Surlykke, 2001] show the potential effectiveness of a hear and avoid collision avoidance system for MAVs.

Search and rescue: Employing aerial robots capable of locating sound sources for search and rescue operations bring forward many important advantages and opportunities. In Chapter 2 it was shown how such aerial robots could provide a method of quickly localizing victims in disaster situations through emergency acoustic sources such as personal alarms or safety whistles, that can operate in night time, through foliage and many adverse weather conditions. Furthermore, a team of MAVs could be used to keep track and coordinate different rescue teams using simple acoustic signalling. Also since sound waves are able to bend around obstacles and travel for long distances, audio based controllers could be used on indoor flying robots to pursue the path of sound waves and quickly reach acoustic targets inside buildings, semi-collapsed structures and caves. A collision resilient flying robot [Briod et al., 2014] equipped with the proposed audio-based direction measurement system could be a promising solution for indoor search and rescue operations.

Aerial surveillance: Many applications could be envisioned for audio based MAVs for surveillance purposes. MAVS could potentially be used to patrol areas of interest for localizing interesting sounds such as alarms, sirens, gun fire and explosions. In protected areas, resting MAVs could launch automatically and quickly localize and reach points of interest upon hearing the sound of security alarms or sirens indicating intrusion. Furthermore, the method presented in Chapter 4 could be employed to allow an MAV, upon locating an acoustic target on the ground, to inform the target's position to other team-mates without the need of a communication network and by simply chirping and circling around the target. In addition, audio based MAVs could potentially join the fight against illegal poaching by patrolling and localizing the sound of hunting rifles in dense forests.

Environmental noise monitoring: Another interesting application could be to use MAVs for performing remote acoustic noise monitoring and inspection. Aerial robots could be used to measure the acoustic noise level and generate acoustic maps in many areas such as construction sites, airports, roads, industrial sites and near wind turbines. Perceiving sounds from the air has the advantage of direct line of sight to the acoustic sources that could improve

the detection and localization as the sound waves are less effected by obstacles.

Study of biological systems: Building small aerial robots that could be socially integrated into groups of flying animals could potentially assist in the study of animal behaviours at the individual and collective level. The ability of these robots to detect, locate and interact through the sense of sound could facilitate the animal-robot interaction. A potential future application that the author envision on this area is to employ audio based MAVs to better study and investigate the purpose of the flight calls of nocturnal migratory birds, where existing research has been limited to far away recordings made from stationary ground stations [Hamilton, 1962, Larkin et al., 2002, Sanders and Mennill, 2014]. In addition using audio-based aerial robots could potentially allow modelling of phonotaxis behaviour among insects,[Doherty, 1985, Hoy, 2014, Hedwig and Robert, 2014] for better understanding of these organisms and obtaining efficient robot controllers.

Mapping and Obstacle avoidance: Detecting and localizing echoes of the emitted sound waves could potentially provide the MAVs with important information about the surrounding environment and the presence of large obstacles for collision avoidance. The work presented by [Dokmanić et al., 2013, Dokmanic et al., 2011] describes a method of determining the basic shape of a room based on a microphone array and from a single sound emission. Implementing a similar method on the proposed audio based system could raise new advantages and applications.

Simple human-robot interaction: An audio-based detection and localization system onboard of aerial robots could provide a simple way of interaction between human operators and flying robots, allowing inexperienced users to operate and control the robots through means of acoustic signals. Experiments performed in Section 2.4.3 illustrate an example of attracting the attention and navigating an aerial robot to a desired point using a simple whistle.

Entertainment: A growing interest in employing swarm of aerial robots for art and entertainment applications have been observed over the recent years [Murphy et al., 2011, Schoch et al., 2014]. Groups of small aerial robots have been demonstrated recently that can generate 3D shapes and animation in the sky [Alonso-Mora et al., 2012]. Following on this line, audio based swarming MAVs could potentially be used to obtain distributed algorithms needed for displaying large objects and animations in the sky and furthermore controlling the swarm to react and dance to sound and music.

5.3 Future directions

While many future directions were suggested in the previous section to reach the potential application propositions, this section presents possible future work mostly aiming at improvements specific to the proposed methods presented in the core of this thesis.

Firstly, improvements could be made by research on the acoustic sensors themselves to

optimize their performance for in-flight measurements. In this work simple, low-cost and commercially available microphone sensors were used that were surely not designed for use on aerial robots. Employing higher quality microphones, or adapting the design of such sensors to obtain the desired sensitivity and frequency response is expected to improve the performance of the proposed methods. In addition, research on wind protectors or anti-vibration mounts for such sensors could potentially allow the reduction of wind and platform-vibrations for obtaining a higher target detection range.

Experiments showed that the self engine noise of robots highly influence the detection range of acoustic targets and increases the rate of false estimates. Research on noise cancelling techniques, both at the mechanical and computational levels could improve the performance of the proposed methods and increase the detection range. Using array of directional microphones pointing away from the self engine, adding an acoustic shield around the robot's engine, and investigating noise cancellation algorithms are some potential future direction on this topic. Figure 3.11.b illustrated a direct relationship between the motor speed and the noise generated by the propeller for a single electrical motor equipped with a propeller. Hence, an active filter for suppressing the engine noise based on the motor input values could potentially be an effective solution.

In this work, methods were mainly based on the time delay of arrival (TDOA) of sound waves between spatially separated microphones. However, other important acoustic information exist that could also be taken into account to increase the robustness and obtain additional information of the target. For example, the perceived sound intensity could be used to obtain an approximate idea about the distance to acoustic targets. Similar to the method explained in Section 2.3.2.1, the Doppler shift in the perceived chirp sound of neighbouring robots could be used to obtain the relative speed with these robots which is an important information for obtaining inter-robot collision avoidance.

Integration of the described audio based systems with other available technologies such as vision and radio waves is an interesting line of research that could rise many additional opportunities. For example, MAVs could compensate the limited field of view of their cameras by actively rotating it towards targets based on the omnidirectional acoustic information. Additionally, integration of acoustic and radio waves is a practical method of obtaining instantaneous relative-range measurements between MAVs. Individuals could emit a radio signal together with the acoustic chirp signal to allow other robots to measure the relative distance based on the time difference between the perceived audio and radio waves.

In Chapter 2, methods for localizing emergency acoustic targets on the ground from a single MAV was described. The solutions could potentially be extended to achieve cooperative localization using a group of MAVs. Measuring the sound direction from multiple MAVs at different points in space, and sharing these information through a communication network, could be used to determine the target's 3D position by simple triangulation. Furthermore, methods for detecting and localizing other type of sound sources such the sound of victims

that are shouting or the sound of a gunfire is of importance.

The proposed relative positioning and localization methods could also be investigated further for measuring and improving the performance of the methods. Finding the minimum rate of the acoustic chirps, i.e. the minimum bearing measurement rate, for performing a specific coordination behaviour, would reduce the power consumption of the system and potentially improve the performance by lowering the number of overlapping chirps. Additionally local behaviours could be investigated to find behaviours that result in increase of precision and a faster convergence in the position estimations. Furthermore, instead of using only linear chirps for the localization, different chirp patterns could potentially be employed to also allow the robots, similar to birds, to communicate simple messages with each-other. In the self localization method presented in Chapter 4, employing multiple beacon robots and obtaining switching protocols for switching the role of robots between the beacon and observer states, could be used to improve localization precision and increase the coverage range.

Bibliography

[Allred et al., 2007] Allred, J., Hasan, A. B., Panichsakul, S., Pisano, W., Gray, P., Huang, J., Han, R., Lawrence, D., and Mohseni, K. (2007). Sensorflock: an airborne wireless sensor network of micro-air vehicles. In *Proceedings of the 5th International Conference on Embedded networked sensor systems*, pages 117–129. ACM.

[Alonso-Mora et al., 2012] Alonso-Mora, J., Schoch, M., Breitenmoser, A., Siegwart, R., and Beardsley, P. (2012). Object and animation display with multiple aerial vehicles. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1078–1083. IEEE.

[Altshuler et al., 2008] Altshuler, Y., Yanovsky, V., Wagner, I. A., and Bruckstein, A. M. (2008). Efficient cooperative search of smart targets using uav swarms. *Robotica*, 26(04):551–557.

[Altuğ et al., 2005] Altuğ, E., Ostrowski, J. P., and Taylor, C. J. (2005). Control of a quadrotor helicopter using dual camera visual feedback. *The International Journal of Robotics Research*, 24(5):329–341.

[Anderson et al., 2008] Anderson, B. D., Fidan, B., Yu, C., and Walle, D. (2008). Uav formation control: theory and application. In *Recent advances in learning and control*, pages 15–33. Springer.

[Andriluka et al., 2010] Andriluka, M., Schnitzspan, P., Meyer, J., Kohlbrecher, S., Petersen, K., Von Stryk, O., Roth, S., and Schiele, B. (2010). Vision based victim detection from unmanned aerial vehicles. In *Proceedings of 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1740–1747. IEEE.

[Artieda et al., 2009] Artieda, J., Sebastian, J. M., Campoy, P., Correa, J. F., Mondragón, I. F., Martínez, C., and Olivares, M. (2009). Visual 3-d slam from uavs. *Journal of Intelligent and Robotic Systems*, 55(4-5):299–321.

[Arulampalam et al., 2002] Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188.

[Asoh et al., 2004] Asoh, H., Asano, F., Yoshimura, T., Yamamoto, K., Motomura, Y., Ichimura, N., Hara, I., and Ogata, J. (2004). An application of a particle filter to bayesian multiple

sound source tracking with audio and video information fusion. In *Procceedings of the 7th International Conference on Information Fusion (IF)*, pages 805–812.

[Basiri et al., 2010] Basiri, M., Bishop, A., and Jensfelt, P. (2010). Distributed control of triangular formations with angle-only constraints. *Systems & Control Letters*, 59(2):147–154.

[Basiri et al., 2013] Basiri, M., Schill, F., Floreano, D., and Lima, P. U. (2013). Audio-based relative positioning system for multiple micro air vehicle systems. In *Proceedings of Robotics: Science and Systems, Berlin, Germany*.

[Basiri et al., 2014a] Basiri, M., Schill, F., Floreano, D., and Lima, P. U. (2014a). Audio-based localization for swarms of micro air vehicles. In *2014 IEEE International Conference on Robotics and Automation (ICRA), Hongkong*.

[Basiri et al., 2012] Basiri, M., Schill, F., Lima, P., and Floreano, D. (2012). Robust acoustic source localization of emergency signals from micro air vehicles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4737 –4742.

[Basiri et al., 2014b] Basiri, M., Schill, F., Lima, P. U., and Floreano, D. (2014b). Localization of emergency acoustic sources by micro aerial vehicles. *Under review in IEEE Transactions on Robotics.*

[Basiri et al., 2015] Basiri, M., Schill, F., Lima, P. U., and Floreano, D. (2015). Onboard audio-based relative positioning system for teams of small aerial robots. *Under submission.*

[Beard et al., 2006] Beard, R. W., McLain, T. W., Nelson, D. B., Kingston, D., and Johanson, D. (2006). Decentralized cooperative aerial surveillance using fixed-wing miniature uavs. *Proceedings of the IEEE*, 94(7):1306–1324.

[Becker et al., 2012] Becker, M., Sampaio, R. C. B., Bouabdallah, S., Perrot, V., and Siegwart, R. (2012). In flight collision avoidance for a mini-uav robot based on onboard sensors. *Mechatronics Lab.*

[Beyeler et al., 2009] Beyeler, A., Zufferey, J.-C., and Floreano, D. (2009). Vision-based control of near-obstacle flight. *Autonomous robots*, 27(3):201–219.

[Blosch et al., 2010] Blosch, M., Weiss, S., Scaramuzza, D., and Siegwart, R. (2010). Vision based mav navigation in unknown and unstructured environments. In *IEEE International Conference on Robotics and automation (ICRA)*.

[Brenneke et al., 2003] Brenneke, C., Wulf, O., and Wagner, B. (2003). Using 3d laser range data for slam in outdoor environments. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS)*, volume 1, pages 188–193. IEEE.

[Briod et al., 2014] Briod, A., Kornatowski, P., Zufferey, J.-C., and Floreano, D. (2014). A collision-resilient flying robot. *Journal of Field Robotics*, 31(4):496–509.

[Brutti et al., 2008a] Brutti, A., Omologo, M., and Svaizer, P. (2008a). Comparison between different sound source localization techniques based on a real data collection. In *IEEE Hands-Free Speech Communication and Microphone Arrays, 2008. HSCMA 2008*, pages 69–72.

[Brutti et al., 2008b] Brutti, A., Omologo, M., and Svaizer, P. (2008b). Localization of multiple speakers based on a two step acoustic map analysis. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4349–4352. IEEE.

[Bryson and Sukkarieh, 2007] Bryson, M. and Sukkarieh, S. (2007). Building a robust implementation of bearing-only inertial slam for a uav. *Journal of Field Robotics*, 24(1-2):113–143.

[Buchner et al., 2005] Buchner, H., Aichner, R., Stenglein, J., Teutsch, H., and Kellennann, W. (2005). Simultaneous localization of multiple sound sources using blind adaptive mimo filtering. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005 (ICASSP'05)*, volume 3, pages iii–97.

[Burgard et al., 2000] Burgard, W., Moors, M., Fox, D., Simmons, R., and Thrun, S. (2000). Collaborative multi-robot exploration. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 476–481. IEEE.

[Byrne et al., 2006] Byrne, J., Cosgrove, M., and Mehra, R. (2006). Stereo based obstacle detection for an unmanned air vehicle. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 2830–2835. IEEE.

[Capus and Brown, 2003] Capus, C. and Brown, K. (2003). Short-time fractional fourier methods for the time-frequency representation of chirp signals. *The Journal of the Acoustical Society of America*, 113(6):3253–3263.

[Carnie et al., 2006] Carnie, R., Walker, R., and Corke, P. (2006). Image processing algorithms for uav "sense and avoid". In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2848 –2853.

[Cole et al., 2008] Cole, D. T., Göktogan, A. H., and Sukkarieh, S. (2008). The demonstration of a cooperative control architecture for uav teams. In *Experimental Robotics*, pages 501–510. Springer.

[Conte and Doherty, 2008] Conte, G. and Doherty, P. (2008). An integrated uav navigation system based on aerial image matching. In *Prpceedings of the 2008 IEEE Aerospace Conference*, pages 1–10. IEEE.

[Corke et al., 2005] Corke, P, Peterson, R., and Rus, D. (2005). Networked robots: Flying robot navigation using a sensor net. In *Robotics Research*, pages 234–243. Springer.

[Das et al., 2002] Das, A. K., Fierro, R., Kumar, V., Ostrowski, J. P., Spletzer, J., and Taylor, C. J. (2002). A vision-based formation control framework. *IEEE Transactions on Robotics and Automation*, 18(5):813–825.

Bibliography

[de Bree et al., 2010] de Bree, H., Wind, I. J., Druyvesteyn, I. E., and te Kulve, H. (2010). Multi purpose acoustic vector sensors for battlefield acoustics. In *Proceedings of the DAMA Conference*.

[DeLima et al., 2006] DeLima, P., York, G., and Pack, D. (2006). Localization of ground targets using a flying sensor network. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, volume 1, page 6 pp.

[DiBiase et al., 2001] DiBiase, J. H., Silverman, H. F., and Brandstein, M. S. (2001). Robust localization in reverberant rooms. In *Microphone Arrays*, pages 157–180. Springer.

[Doherty, 1985] Doherty, J. A. (1985). Trade-off phenomena in calling song recognition and phonotaxis in the cricket, gryllus bimaculatus (orthoptera, gryllidae). *Journal of Comparative Physiology A*, 156(6):787–801.

[Dokmanic et al., 2011] Dokmanic, I., Lu, Y. M., and Vetterli, M. (2011). Can one hear the shape of a room: The 2-d polygonal case. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 321–324. IEEE.

[Dokmanić et al., 2013] Dokmanić, I., Parhizkar, R., Walther, A., Lu, Y. M., and Vetterli, M. (2013). Acoustic echoes reveal room shape. *Proceedings of the National Academy of Sciences*, 110(30):12186–12191.

[Doucet et al., 2001] Doucet, A., De Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer Verlag.

[Doyle et al., 2013] Doyle, C. E., Bird, J. J., Isom, T. A., Kallman, J. C., Bareiss, D. F., Dunlop, D. J., King, R. J., Abbott, J. J., and Minor, M. A. (2013). An avian-inspired passive mechanism for quadrotor perching. *IEEE/ASME Transactions on Mechatronics*, 18(2):506–517.

[Durak and Aldirmaz, 2010] Durak, L. and Aldirmaz, S. (2010). Adaptive fractional fourier domain filtering. *Signal Processing*, 90(4):1188–1196.

[Erden et al., 1999] Erden, M. F., Kutay, M. A., and Ozaktas, H. M. (1999). Repeated filtering in consecutive fractional fourier domains and its application to signal restoration. *IEEE transactions on signal processing*, 47(5):1458–1462.

[Farnsworth, 2005] Farnsworth, A. (2005). Flight calls and their value for future ornithological studies and conservation research. *The Auk*, 122(3):733–746.

[Ferguson, 1999] Ferguson, B. (1999). Time-delay estimation techniques applied to the acoustic detection of jet aircraft transits. *The Journal of the Acoustical Society of America*, 106:255.

[Frew et al., 2007] Frew, E. W., Lawrence, D. A., Dixon, C., Elston, J., and Pisano, W. J. (2007). Lyapunov guidance vector fields for unmanned aircraft applications. In *American Control Conference*, pages 371–376. IEEE.

[Garcia et al., 1996] Garcia, J., Mas, D., and Dorsch, R. G. (1996). Fractional-fourier-transform calculation through the fast-fourier-transform algorithm. *Applied optics*, 35(35):7013–7018.

[Gaszczak et al., 2011] Gaszczak, A., Breckon, T. P., and Han, J. (2011). Real-time people and vehicle detection from uav imagery. In *IS&T/SPIE Electronic Imaging*, pages 78780B–78780B. International Society for Optics and Photonics.

[Gibson et al., 2010] Gibson, G., Warren, B., and Russell, I. (2010). Humming in tune: sex and species recognition by mosquitoes on the wing. *JARO-Journal of the Association for Research in Otolaryngology*, 11(4):527–540.

[Gilks et al., 1996] Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). *Markov chain Monte Carlo in practice*. Chapman & Hall/CRC.

[Goktouan et al., 2010] Goktouan, A. H., Sukkarieh, S., Bryson, M., Randle, J., Lupton, T., and Hung, C. (2010). A rotary-wing unmanned air vehicle for aquatic weed surveillance and management. In *Selected papers from the 2nd International Symposium on UAVs, Reno, Nevada, USA June 8–10, 2009*, pages 467–484. Springer.

[Goodrich et al., 2008] Goodrich, M. A., Morse, B. S., Gerhardt, D., Cooper, J. L., Quigley, M., Adams, J. A., and Humphrey, C. (2008). Supporting wilderness search and rescue using a camera-equipped mini uav. *Journal of Field Robotics*, 25(1-2):89–110.

[Griffiths et al., 2006] Griffiths, S., Saunders, J., Curtis, A., Barber, B., McLain, T., and Beard, R. (2006). Maximizing miniature aerial vehicles. *IEEE Robotics & Automation Magazine*, 13(3):34–43.

[Grocholsky et al., 2006] Grocholsky, B., Keller, J., Kumar, V., and Pappas, G. (2006). Cooperative air and ground surveillance. *Robotics & Automation Magazine, IEEE*, 13(3):16–25.

[Hamilton, 1962] Hamilton, W. J. (1962). Evidence concerning the function of nocturnal call notes of migratory birds. *The Condor*, pages 390–401.

[Hauert et al., 2011] Hauert, S., Leven, S., Varga, M., Ruini, F., Cangelosi, A., Zufferey, J.-C., and Floreano, D. (2011). Reynolds flocking in reality with fixed-wing robots: communication range vs. maximum turning rate. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5015–5020. IEEE.

[Hauert et al., 2010] Hauert, S., Leven, S., Zufferey, J., and Floreano, D. (2010). Communication-based swarming for flying robots. In *Proceedings of International Conference on Robotics and Automation, Workshop on Network Science and Systems, Anchorage, Alaska*.

[Hebert, 2000] Hebert, M. (2000). Active and passive range sensing for robotics. In *Proceedings of the IEEE International Conference on Robotics and Automation ICRA*, volume 1, pages 102–110 vol.1.

[Hedwig and Robert, 2014] Hedwig, B. and Robert, D. (2014). Auditory parasitoid flies exploiting acoustic communication of insects. In *Insect Hearing and Acoustic Communication*, pages 45–63. Springer.

[Hoy, 2014] Hoy, R. R. (2014). Hearing in insects: The why, when, and how. In *Perspectives on Auditory Research*, pages 287–298. Springer.

[Hu et al., 2011] Hu, J.-S., Tsai, C.-M., Chan, C.-Y., and Chang, Y.-J. (2011). Geometrical arrangement of microphone array for accuracy enhancement in sound source localization. In *Proceedings of IEEE 8th Asian Control Conference (ASCC), Kaohsiung, Taiwan*, pages 299–304.

[Humphreys, 2012] Humphreys, T. (2012). Statement on the vulnerability of civil unmanned aerial vehicles and other systems to civil gps spoofing. *University of Texas at Austin (July 18, 2012)*.

[James et al., 2001] James, C. et al. (2001). Vulnerability assessment of the transportation infrastructure relying on the global positioning system. *Volpe National Transportation Systems Center, US Department of Transportation, Tech. Rep.*

[Jensfelt and Christensen, 2001] Jensfelt, P. and Christensen, H. I. (2001). Pose tracking using laser scanning and minimalistic environmental models. *IEEE Transactions on Robotics and Automation*, 17(2):138–147.

[Katrasnik et al., 2008] Katrasnik, J., Pernus, F., and Likar, B. (2008). New robot for power line inspection. In *Robotics, Automation and Mechatronics, 2008 IEEE Conference on*, pages 1195–1200. IEEE.

[Kemppainen et al., 2006] Kemppainen, A., Haverinen, J., and Röning, J. (2006). An infrared location system for relative pose estimation of robots. In Zielińska, T. and Zieliński, C., editors, *Romansy 16*, volume 487 of *CISM Courses and Lectures*, pages 379–386. Springer Vienna.

[Kernbach, 2013] Kernbach, S. (2013). *Handbook of collective robotics: fundamentals and challenges*. CRC Press.

[Knapp and Carter, 1976] Knapp, C. and Carter, G. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24(4):320–327.

[Kottege and Zimmer, 2007] Kottege, N. and Zimmer, U. (2007). Relative localisation for auv swarms. In *Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies, 2007.*, pages 588–593. IEEE.

[Kovacina et al., 2002] Kovacina, M. A., Palmer, D., Yang, G., and Vaidyanathan, R. (2002). Multi-agent control algorithms for chemical cloud detection and mapping using unmanned air vehicles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 2782–2788. IEEE.

[Kownacki, 2011] Kownacki, C. (2011). Obstacle avoidance strategy for micro aerial vehicle. In *Advances in Aerospace Guidance, Navigation and Control*, pages 117–135. Springer.

[Kushleyev et al., 2013] Kushleyev, A., Mellinger, D., Powers, C., and Kumar, V. (2013). Towards a swarm of agile micro quadrotors. *Autonomous Robots*, 35(4):287–300.

[Kutay et al., 1997] Kutay, M. A., Ozaktas, H. M., Ankan, O., and Onural, L. (1997). Optimal filtering in fractional fourier domains. *IEEE Transactions on Signal Processing*, 45(5):1129–1143.

[Lanzisera et al., 2006] Lanzisera, S., Lin, D. T., and Pister, K. S. (2006). Rf time of flight ranging for wireless sensor network localization. In *International Workshop on Intelligent Solutions in Embedded Systems*, pages 1–12. IEEE.

[Larkin et al., 2002] Larkin, R. P., Evans, W. R., and Diehl, R. H. (2002). Nocturnal flight calls of dickcissels and doppler radar echoes over south texas in spring. *Journal of Field Ornithology*, 73(1):2–8.

[Lawson and Hanson, 1995] Lawson, C. and Hanson, R. (1995). *Solving least squares problems*, volume 15. Society for Industrial Mathematics.

[Leven et al., 2007] Leven, S., Zufferey, J., and Floreano, D. (2007). A simple and robust fixed-wing platform for outdoor flying robot experiments. In *International symposium on flying insects and robots*, pages 69–70.

[Lindsey et al., 2012] Lindsey, Q., Mellinger, D., and Kumar, V. (2012). Construction with quadrotor teams. *Autonomous Robots*, 33(3):323–336.

[Lupashin et al., 2010] Lupashin, S., Schollig, A., Sherback, M., and D'Andrea, R. (2010). A simple learning strategy for high-speed quadrocopter multi-flips. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1642–1648. IEEE.

[Mataric, 1993] Mataric, M. J. (1993). Designing emergent behaviors: From local interactions to collective intelligence. In *Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 432–441.

[Mataric, 1997] Mataric, M. J. (1997). Behaviour-based control: examples from navigation, learning, and group behaviour. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2-3):323–336.

[Matsusaka et al., 1999] Matsusaka, Y., Tojo, T., Kubota, S., Furukawa, K., Tamiya, D., Hayata, K., Nakano, Y., and Kobayashi, T. (1999). Multi-person conversation via multi-modal interface-a robot who communicate with multi-user. In *Proceedings of Eurospeech*, pages 1723–1726.

[McLurkin and Smith, 2007] McLurkin, J. and Smith, J. (2007). Distributed algorithms for dispersion in indoor environments using a swarm of autonomous mobile robots. In *Distributed Autonomous Robotic Systems 6*, pages 399–408. Springer.

Bibliography

[Mejias et al., 2010] Mejias, L., McNamara, S., Lai, J., and Ford, J. (2010). Vision-based detection and tracking of aerial targets for uav collision avoidance. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 87–92. IEEE.

[Melhuish et al., 2002] Melhuish, C., Welsby, J., and Greenway, P. (2002). Gradient ascent with a group of minimalist real robots: Implementing secondary swarming. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, volume 2, pages 509–514.

[Mellinger et al., 2010] Mellinger, D., Shomin, M., and Kumar, V. (2010). Control of quadrotors for robust perching and landing. In *Proceedings of the International Powered Lift Conference*.

[Merino et al., 2006] Merino, L., Caballero, F., Martínez-de Dios, J. R., Ferruz, J., and Ollero, A. (2006). A cooperative perception system for multiple uavs: Application to automatic detection of forest fires. *Journal of Field Robotics*, 23(3-4):165–184.

[Michael et al., 2011] Michael, N., Fink, J., and Kumar, V. (2011). Cooperative manipulation and transportation with aerial robots. *Autonomous Robots*, 30(1):73–86.

[Miller and Surlykke, 2001] Miller, L. A. and Surlykke, A. (2001). How some insects detect and avoid being eaten by bats: Tactics and countertactics of prey and predator. *Bioscience*, 51(7):570–581.

[Miro, 2007] Miro, X. A. (2007). *Robust speaker diarization for meetings*. Universitat Politècnica de Catalunya.

[Moshtagh et al., 2009] Moshtagh, N., Michael, N., Jadbabaie, A., and Daniilidis, K. (2009). Vision-based, distributed control laws for motion coordination of nonholonomic robots. *IEEE Transactions on Robotics*, 25(4):851 –860.

[Muller et al., 2014] Muller, J., Ruiz, A., and Wieser, I. (2014). Safe amp; sound: A robust collision avoidance layer for aerial robots based on acoustic sensors. In *Position, Location and Navigation Symposium - PLANS 2014, 2014 IEEE/ION*, pages 1197–1202.

[Muller and Robert, 2001] Muller, P. and Robert, D. (2001). A shot in the dark: the silent quest of a free-flying phonotactic fly. *Journal of Experimental Biology*, 204(6):1039–1052.

[Murphy et al., 2011] Murphy, R., Shell, D., Guerin, A., Duncan, B., Fine, B., Pratt, K., and Zourntos, T. (2011). A midsummer night's dream (with flying robots). *Autonomous Robots*, 30(2):143–156.

[Nakadai et al., 2000] Nakadai, K., Lourens, T., Okuno, H., and Kitano, H. (2000). Active audition for humanoid. In *AAAI-2000*, pages 832–839. MIT Press.

[Namias, 1980] Namias, V. (1980). The fractional order fourier transform and its application to quantum mechanics. *IMA Journal of Applied Mathematics*, 25(3):241–265.

[Nelson et al., 2007] Nelson, D. R., Barber, D. B., McLain, T. W., and Beard, R. W. (2007). Vector field path following for miniature air vehicles. *IEEE Transactions on Robotics*, 23(3):519–529.

[Nüchter and Hertzberg, 2008] Nüchter, A. and Hertzberg, J. (2008). Towards semantic maps for mobile robots. *Robotics and Autonomous Systems*, 56(11):915–926.

[Okuno et al., 2002] Okuno, H., Nakadai, K., and Kitano, H. (2002). Social interaction of humanoid robot based on audio-visual tracking. *Developments in Applied Artificial Intelligence*, pages 140–173.

[Oyekan and Huosheng, 2009] Oyekan, J. and Huosheng, H. (2009). Toward bacterial swarm for environmental monitoring. In *IEEE International Conference on Automation and Logistics, ICAL'09.*, pages 399–404. IEEE.

[Ozaktas et al., 1994] Ozaktas, H., Barshan, B., Mendlovic, D., and Onural, L. (1994). Convolution, filtering, and multiplexing in fractional fourier domains and their relation to chirp and wavelet transforms. *Journal of the Optical Society of America*, 11(2):547–559.

[Pack et al., 2006] Pack, D., York, G., and Fierro, R. (2006). Information-based cooperative control for multiple unmanned aerial vehicles. In *Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, ICNSC*, pages 446 –450.

[Parunak et al., 2003] Parunak, H. V. D., Brueckner, S., and Odell, J. J. (2003). Swarming coordination of multiple uav's for collaborative sensing. In *Proceedings of the Second AIAA" Unmanned Unlimited" Systems, Technologies, and Operations Conference.*

[Perez-Lorenzo et al., 2012] Perez-Lorenzo, J., Viciana-Abad, R., Reche-Lopez, P., Rivas, F., and Escolano, J. (2012). Evaluation of generalized cross-correlation methods for direction of arrival estimation using two microphones in real environments. *Applied Acoustics*, 73(8):698–712.

[Pinker and Smith, 1999] Pinker, A. and Smith, C. (1999). Vulnerability of the gps signal to jamming. *GPS Solutions*, 3(2):19–27.

[Pollack, 2000] Pollack, G. (2000). Who, what, where? recognition and localization of acoustic signals by insects. *Current opinion in neurobiology*, 10(6):763–767.

[Populin, 2006] Populin, L. C. (2006). Monkey sound localization: head-restrained versus head-unrestrained orienting. *The Journal of neuroscience*, 26(38):9820–9832.

[Pugh and Martinoli, 2006] Pugh, J. and Martinoli, A. (2006). Relative localization and communication module for small-scale multi-robot systems. In *Proceedings IEEE International Conference on Robotics and Automation, ICRA*, pages 188 –193.

[Pugh et al., 2009] Pugh, J., Raemy, X., Favre, C., Falconi, R., and Martinoli, A. (2009). A fast onboard relative positioning module for multirobot systems. *IEEE/ASME Transactions on Mechatronics*, 14(2):151–162.

[Puri, 2005] Puri, A. (2005). A survey of unmanned aerial vehicles (uav) for traffic surveillance. *Department of computer science and engineering, University of South Florida.*

Bibliography

[Reilly et al., 2010] Reilly, V., Solmaz, B., and Shah, M. (2010). Geometric constraints for human detection in aerial imagery. In *Computer Vision–ECCV 2010*, pages 252–265. Springer.

[Reynolds, 1987] Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.*, 21(4):25–34.

[Ritz et al., 2012] Ritz, R., Muller, M., Hehn, M., and D'Andrea, R. (2012). Cooperative quadrocopter ball throwing and catching. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4972–4978. IEEE.

[Roberts et al., 2009] Roberts, J., Stirling, T., Zufferey, J.-C., and Floreano, D. (2009). 2.5d infrared range and bearing system for collective robotics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pages 3659 –3664.

[Roberts et al., 2008] Roberts, J. F., Zufferey, J.-C., and Floreano, D. (2008). Energy management for indoor hovering robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS)*, pages 1242–1247. IEEE.

[Ruffier et al., 2011] Ruffier, F., Benacchio, S., Expert, F., and Ogam, E. (2011). A tiny directional sound sensor inspired by crickets designed for micro-air vehicles. In *Sensors, 2011 IEEE*, pages 970–973. IEEE.

[Ruini and Cangelosi, 2009] Ruini, F. and Cangelosi, A. (2009). Extending the evolutionary robotics approach to flying machines: An application to mav teams. *Neural Networks*, 22(5):812–821.

[Sa and Corke, 2014] Sa, I. and Corke, P. (2014). Vertical infrastructure inspection using a quadcopter and shared autonomy control. In *Field and Service Robotics*, pages 219–232. Springer.

[Sanders and Mennill, 2014] Sanders, C. E. and Mennill, D. J. (2014). Acoustic monitoring of nocturnally migrating birds accurately assesses the timing and magnitude of migration through the great lakes. *The Condor*, 116(3):371–383.

[Saunders et al., 2005] Saunders, J. B., Call, B., Curtis, A., Beard, R. W., and McLain, T. W. (2005). Static and dynamic obstacle avoidance in miniature air vehicles. *AIAA Infotech at Aerospace*.

[Sauter et al., 2005] Sauter, J. A., Matthews, R., Van Dyke Parunak, H., and Brueckner, S. A. (2005). Performance of digital pheromones for swarming vehicle control. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 903–910. ACM.

[Scherer et al., 2007] Scherer, S., Singh, S., Chamberlain, L., and Saripalli, S. (2007). Flying fast and low among obstacles. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2023–2029. IEEE.

[Schoch et al., 2014] Schoch, M., Alonso-Mora, J., Siegwart, R., and Beardsley, P. (2014). Viewpoint and trajectory optimization for animation display with aerial vehicles. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 4711–4716. IEEE.

[Shim et al., 2003] Shim, D. H., Kim, H. J., and Sastry, S. (2003). Decentralized nonlinear model predictive control of multiple flying robots. In *Proceedings of the 42nd IEEE conference on Decision and control*, volume 4, pages 3621–3626. IEEE.

[Siegwart and Nourbakhsh, 2004] Siegwart, R. and Nourbakhsh, I. (2004). *Introduction to autonomous mobile robots*. MIT press.

[Simmons et al., 2000] Simmons, R., Apfelbaum, D., Burgard, W., Fox, D., Moors, M., Thrun, S., and Younes, H. (2000). Coordination for multi-robot exploration and mapping. In *AAAI/IAAI*, pages 852–858.

[Stern et al., 1988] Stern, R., Zeiberg, A., and Trahiotis, C. (1988). Lateralization of complex binaural stimuli: A weighted-image model. *The Journal of the Acoustical Society of America*, 84:156.

[Stirling et al., 2012] Stirling, T., Roberts, J., Zufferey, J.-C., and Floreano, D. (2012). Indoor navigation with a swarm of flying robots. In *Proceedings of the 2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4641–4647.

[Templeton et al., 2007] Templeton, T., Shim, D. H., Geyer, C., and Sastry, S. S. (2007). Autonomous vision-based landing and terrain mapping using an mpc-controlled unmanned rotorcraft. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1349–1356. IEEE.

[Utt et al., 2005] Utt, J., McCalmont, J., and Deschenes, M. (2005). Development of a sense and avoid system. *AIAA Infotech at Aerospace*.

[Valenti et al., 2007] Valenti, M., Bethke, B., Dale, D., Frank, A., McGrew, J., Ahrens, S., How, J. P., and Vian, J. (2007). The mit indoor multi-vehicle flight testbed. In *IEEE International Conference on Robotics and Automation*, pages 2758–2759. IEEE.

[Valin et al., 2007] Valin, J., Michaud, F., and Rouat, J. (2007). Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. *Robotics and Autonomous Systems*, 55(3):216–228.

[Valin et al., 2003] Valin, J., Michaud, F., Rouat, J., and Létourneau, D. (2003). Robust sound source localization using a microphone array on a mobile robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems , IROS*, volume 2, pages 1228–1233. IEEE.

[Viquerat et al., 2008] Viquerat, A., Blackhall, L., Reid, A., Sukkarieh, S., and Brooker, G. (2008). Reactive collision avoidance for unmanned aerial vehicles using doppler radar. In *Field and Service Robotics*, pages 245–254. Springer.

Bibliography

[Ward et al., 2003] Ward, D., Lehmann, E., and Williamson, R. (2003). Particle filtering algorithms for tracking an acoustic source in a reverberant environment. *Speech and Audio Processing*, 11(6):826–836.

[Weiss et al., 2011] Weiss, S., Scaramuzza, D., and Siegwart, R. (2011). Monocular-slam–based navigation for autonomous micro helicopters in gps-denied environments. *Journal of Field Robotics*, 28(6):854–874.

[Welch and Bishop, 1995] Welch, G. and Bishop, G. (1995). An introduction to the kalman filter.

[Yang et al., 2005] Yang, Y., Minai, A., and Polycarpou, M. M. (2005). Evidential map-building approaches for multi-uav cooperative search. In *Proceedings of the American Control Conference*, volume 1, page 116.

Meysam Basiri
Date of birth : 1982-09-11

Contact

E-mail : meysam.basiri @epfl.ch,   basiri@kth.se
Url: people.epfl.ch/basiri
Tel : +41787299993

Address

EPFL STI IMT LIS ELE134, Station 11, CH-1015 Lausanne, Switzerland

Research Interests

Aerial robotics	Localization & Navigation
Distributed intelligent systems	Control systems
Signal processing	Acoustics / Computer vision

Education

2010 to 2015	**Ph.D in Robotics. Topic: "Audio-based Swarming Micro Aerial Vehicles"**
	* Laboratory of Intelligence Systems, École Polytechnique Fédérale de Lausanne, EPFL, Switzerland.
	* Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal.
2007 to 2010	**Master in Systems, Control and Robotics** KTH, Royal Institute of Technology, Stockholm, Sweden
2001 to 2007	**Bachelor in Electrical Engineering - "Control"** IUST, Iran University of Science and Technology, Tehran, Iran
1999 to 2000	**Stage 1-Bsc(Hons) in Electrical and Electronics Engineering** University of Derby, Derby, United Kingdom
1998 to 1999	**A-level "Mathematics, Physics"** Broxtowe College, Nottingham, United Kingdom

Publications

- ✓ M. Basiri, F. S. Schill, D. Floreano and P. Lima. *Localization of emergency sound sources by micro aerial vehicles,* under review by IEEE Transactions on Robotics Journal.
- ✓ M. Basiri, F. S. Schill, D. Floreano and P. Lima. *Audio-based localization for swarms of micro air vehicles*. IEEE International Conference on Robotics and Automation (ICRA 2014), Hong Kong, China,
- ✓ M. Basiri, F. Schill, D. Floreano and P. Lima. *Audio-based relative positioning system for multiple micro air vehicle systems*. Robotics: Science and Systems RSS2013, Berlin, Germany, 2013.
- ✓ M. Basiri, F. S. Schill, P. Lima U. and D. Floreano. *Robust acoustic source localization of emergency signals from micro air vehicles*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura, Algarve, Portugal, 2012.
- ✓ Basiri, M. & Bishop, A. & Jensfelt, P. *Distributed control of triangular formations with angle-only constraints*, Systems & Control Letters Journal.
- ✓ Basiri, M. & Bishop, A. & Jensfelt, P. *Distributed control of sensor formations with angle-only constraints.* ISSNIP 2009 ,Melbourne, Australia, to appear on IEEE Xplore
- ✓ Basiri, M. & Bishop, A. & Jensfelt, P. *Distributed formation control with angle-only constraints*, Swedish Workshop on Autonomous Robots, 2009
- ✓ Basiri, M. & Bishop, A. & Jensfelt, P, *Bearing-only triangular formation control on the plane and the sphere*, 18th Mediterranean Conference on Control & Automation, June 23-25, 2010

Awards and Professional Memberships

- ✓ IROS 2012 Best application paper award.
- ✓ IROS 2012 Best student paper finalist award
- ✓ Academic Excellent Award in year 2006
- ✓ Fourth ARTIST2 graduate member on Embedded Control Systems (May 26-30, 2008, KTH)
- ✓ IUST Robotic group member

Projects and Research

Major Qualifying Projects

✓ Audio-based swarming micro air vehicles (PhD Thesis)	2014
✓ Distributed formation control of multiple agents with angle-only constraints (MSc Thesis)	2009
✓ Multi Robot localization based on active sonar system (BSc Thesis)	2007

Research Projects

- ✓ Acoustic target localization for aerial robots, LIS, EPFL.
- ✓ Audio-based relative positioning system for teams of Micro air vehicles, LIS, EPFL.
- ✓ Emergency wireless sensor network using swarms of aerial robots, SMAVNET project, LIS, EPFL.
- ✓ Portable acoustic localization module for localizing aerial robots using their engine sounds. LIS, EPFL.
- ✓ Bearing only formation control for multi MAV systems. CSC, KTH.
- ✓ Control of a 2-DOF Inverted Pendulum , Automatic Control, KTH
- ✓ Mobile rescuing robot, Robotics and Autonomous Systems, CSC, KTH.
- ✓ 2-DOF Robot Arm Control, Automatic Control Lab,S3, KTH
- ✓ Table detection using stereovision, CogX project, CSC, KTH
- ✓ Autonomous motion control of a mobile robot in corridors, CSC, KTH
- ✓ Vision control of a ball grabbing mobile robot.
- ✓ Mobile Robot localization and tracking using "Extended Kalman Filtering" and "Particle Filtering"
- ✓ Active beacon ultrasonic system for tracking and guidance of a mobile robot, IUST

Other Projects

- ✓ Development of embedded autopilot and audio-based estimator/controllers for micro aerial robots.
- ✓ Modeling and control design of a double tank process, Hybrid and Embedded control
- ✓ Temperature monitoring and control of a green house
- ✓ Multi sensor data reading and communication using Microcontrollers
- ✓ Ultrasonic range finder and IR transceiver
- ✓ PID speed controller for a dc motor using PIC microcontroller

Background Courses Include

Distributed Intelligent Systems	Mobile Robots	Applied Estimation
Nonlinear Control	Microcontrollers	Computer Vision
Hybrid & Embedded Control	Scientific programming	Artificial Intelligence
Modern Control	Mechatronics	Instrumentation
Digital Control	Image Processing	Project management
Industrial Control	Modeling of Dynamic systems	Robotics and Autonomous Systems
Signal Processing	Electronics	

Computer Skills

- ✓ C, C++, python languages
- ✓ Matlab and Simulink
- ✓ LabView
- ✓ Microcontroller programming (AVR studio, CodeVision AVR, PicBasic, MPLAB)
- ✓ Webots
- ✓ Mathematica
- ✓ Eagle PCB circuit design
- ✓ Proteus
- ✓ PIPE and HyVisual
- ✓ Web design and Animation programs (Adobe Dreamweaver, Adobe Flash, Adobe Premier, Studio Max)
- ✓ LaTeX, Word, Powerpoint, Excell , Illustrator

Work & Teaching Experience

2010-2014   **École polytechnique fédérale de Lausanne, EPFL**   Lausanne, Switzerland

- ▪ Teaching assistant in bio-inspired artificial intelligence courses
- ▪ Teaching assistant in mobile robotics courses
- ▪ Main supervisor for 7 student projects at masters level
- ▪ Gave demonstrations & presentations of research outcomes of LIS

Summer 2008   **Royal Institute of Technology, Center of Autonomous Systems** Stockholm, Sweden, Research on "Surface detection and classification using stereovision" CogX project

2006 -2007   **Iran University of Science and Technology**   Tehran, Iran
Research Assistant at Dr Shahri Research Laboratory

2005 - 2006   **Iran University of Science and Technology**   Tehran, Iran
Teaching Assistant in "Logic Circuits" and "Instrumentation" courses

2005   **Tozin Taraz Aria Instrumentation CO**   (Internship)   Tehran, Iran
Microcontroller programming and Hardware setup of digital scales

2003 - 2007   **Matin Rayane Computer Corp**   (Part time)   Tehran, Iran
Programming, Multimedia and Web design

Languages

Persian (native)   **English** (fluent)   **French/Portuguese** (beginner)

Hobbies

Piloting quad-rotors and model airplanes, Hiking, Traveling , Football, Programming