Design of a Radio Direction Finder for Search and Rescue Operations: Estimation, Sonification, and Virtual Prototyping

David Poirier-Quinot

LIMSI
Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur

ETIS
Equipes Traitement de l'Information et Systèmes

Airbus Defense & Space

Design of a Radio Direction Finder for Search and Rescue Operations

Estimation, Sonification, and Virtual Prototyping

A Thesis to obtain the title of
Ph.D. of Science
Pierre and Marie Curie University

David Poirier-Quinot

Thesis Advisor: Brian F.G. Katz

Jury:

Reviewers:	Peres S. Camille	-	Texas A&M Health Science Center
	Kuhlen Torsten W.	-	RWTH Aachen University
Advisor:	Katz Brian F.G.	-	LIMSI, CNRS
President:	Adam Olivier	-	LAM, UPMC
	Fijalkow Inbar	-	ETIS, ENSEA
	Girardeau Laurent	-	Airbus Defense & Space

Specialty: Computer Sciences, Acoustics and Electronics

April 2015

Abstract

This research investigates the design of a radio Direction Finder (DF) for rescue operation using victims' cellphone as localization beacons. The conception is focused on an audio interface, using sound to progressively guide rescuers towards their target. The thesis' ambition is to exploit the natural mechanisms of human hearing to improve the global performance of the search process rather than to develop new Direction-Of-Arrival (DOA) estimation techniques.

Classical DOA estimation techniques are introduced along with a range of tools to assess their efficiency. Based on these tools, a case study is proposed regarding the performance that might be expected from a lightweight DF design tailored to portable operation. It is shown that the performance of high-resolution techniques usually implemented for DOA estimation are seriously impacted by any size-constraint applied on the DF, particularly in multi-path propagation conditions.

Subsequently, a review of interactive parameter mapping sonification is proposed. Various sonification paradigms are designed and assessed regarding their capacity to convey information related to different levels of DF outputs. Listening tests are conducted suggesting that trained subjects are capable of monitoring multiple audio streams and gather information from complex sounds. Said tests also indicate the need for a DF sonification that perceptively orders the presented information, for beginners to be able to effortlessly focus on the most important data only. Careful attention is given to sound aesthetic and how it impacts operators' acceptance and trust in the DF, particularly regarding the perception of measurement noise during the navigation.

Finally, a virtual prototype is implemented that recreates DF-based navigation in a virtual environment to evaluate the proposed sonification mappings. In the meantime, a physical prototype is developed to assess the ecological validity of the virtual evaluations. Said prototype exploits a software defined radio architecture for rapid iteration through design implementations. The overall performance evaluation study is conducted in consultation with rescue services representatives and compared with their current search solutions.

It is shown that, in this context, simple DF designs based on the parallel sonification of the output signal of several antennas may produce navigation performance comparable to these of more complex designs based on high-resolution methods. As the task objective is to *progressively* localize a target, the system's cornerstone appears to be the robustness and consistency of its estimations rather than its punctual accuracy. Involving operators in the estimation allows avoiding critical situations where one feels helpless when faced with an autonomous system producing non-sensical estimations. Virtual prototyping proved to be a sensible and efficient method to support this study, allowing for fast iterations through sonifications and DF designs implementations.

Learning is a basic skill, to be honed and cherished, yet innate.
It's teaching that comes hard, particularly teaching oneself.

Declaration

Some of the material gathered in this thesis has previously been made public. Parts of Chapter 5 appeared in a paper published in the proceedings of the 4$^{\text{th}}$ International Congress on Ultra Modern Telecommunications and Control Systems (2012) [1]. Parts of Chapter 12 and Appendix A were published in the proceedings of the 19$^{\text{th}}$ and 20$^{\text{th}}$ International Conference on Auditory Display (2014) [2,3] and accepted for publication in the proceedings of the IEEE VR Conference[1] (2015). Finally, parts of Chapter 9 have been submitted to the International Journal of Human-Computer Studies[2] (2015).

Some of the tools exploited in this thesis are based on previously existing implementations or have been developed in collaboration with other researchers. Julian Adenauer and Jorge Gascon first proposed the concept of what became the BlenderVR software[3], mainly developed by Damien Touraine and Dalai Felinto assisted by Brian F.G. Katz, Laurent Pointal and myself. The IlmProp software[4] was developed by scholars from the Ilmenau Technische Universität, kindly introduced to me by Prof. J.P. C. Lustosa da Costa at the ICUMT Conference in 2012. Other softwares and packages (Matlab, WINNER II, Blender, Max/MSP, Pd, SPAT, GNU Radio etc.) were implemented by various developers, industries and open source communities. My dearest thanks to them all for sharing (or at the very least selling) their work.

Finally, most of the work reported here was supported by a grant from the French government through the ANRT in the context of a CIFRE industry-linked research funding program with Airbus Defense & Space (ID 2011/1309).

Except where otherwise stated, I declare that this thesis is my own original work.

David Poirier-Quinot
April the 18$^{\text{th}}$, 2015

[1] D. Poirier-Quinot and B. F.G. Katz, "BlenderVR: Open-source framework for interactive and immersive VR".

[2] D. Poirier-Quinot, G. Parseihian, and B. F.G. Katz, "Reduction of perceived instabilities in Parameter Mapping Sonification: application to the real-time exploration of a noisy stream of data".

[3] BlenderVR: authoring tool and game engine for Virtual Reality, `http://blendervr.limsi.fr`.

[4] IlmProp: open source raytracing model, `http://www.tu-ilmenau.de/ilmprop`.

Acknowledgements

So much person to thank for these last years, so few lines to do so. I do hope they won't take offense of hereafter "structured thanks" (e.g. per establishment, nothing hierarchical I swear). Structure (regardless of its quality or meaning) somewhat becomes a habit after puzzling a would-be-decent thesis out of three years of what looks like random wanderings from afar.

My first thanks go to those who conducted this research project by my side. Brian K., as my thesis advisor and may I say friend, educated me to Science and the thorough and suspicious mind that goes with it. Laurent G., probably without knowing it (as all good mentors do) showed me how to balance reflection and recklessness to achieve virtually everything, and Inbar F. who ever willingly supported my would-be-scientific ambitions on would-be-relevant research topics. I am forever grateful to Camille P., Torsten K. and Olivier A. for reviewing my thesis, presiding my defense jury and all along nourishing reflections on my contributions to improve the quality of this manuscript.

Second come LIMSI's scholars, from which I learnt much, with which I endured a lot. Albert, Brian, I'll forever relish these talks we had and will have, making everyday's mood around a cup of morning coffee. Christophe, my thanks for your calm and quiescent ways, ever at odds (and all the more welcome) with my overactivity (never hyper, I swear) at work. Gone while I was still a young graduate, I'd like to thank my now pears Ph.D fellows for illustrating the simple yet dubious fact that this tunnel did end somewhere (well, that at the very least digging or C-4ing was an option), from Gaëtan the Gallant (trying to break a long tradition of undeserved nicknames) to Marc R. the "guy who did it before, faster, better". David D., for his nonchalance that went astray with his coding skills, Tiffany B. for resembling so much my sister (though I never told her in so many words), Paul L. and the soothing way he seemed to backpack through his thesis, Matthieu A. for his brief yet cheerful existence (did I say existence?) in Brian's team, and finally Lionel F. so perfect an illustration of the iceberg theorem, with which I wish I'll have moments in the future to investigate the underwater portion.

My dearest thoughts to my daily coworkers, some of them endeavoring to write their own manuscript at this very minute. Marc E. and his knack for asking the one question when I talked nonsense, Samuel D. and his infinite patience when listening to my endless comments over his work, Olivier P. and the shared musical moments, Justin M. and his ever present readiness to brainstorm, Bart P. for reminding me that Science encompasses more than trendy things and that History may be worth opening a few books here and there, and Areti A. for her sharp scientific mind and direct-slash-sincere ways. I shall always hold fond memories of all the people I worked with in the other research teams of Orsay: Damien T. my coding mentor and support throughout

countless miseries, Jean-Marc V. the humble Gyro Gearloose, Nicolas F. the calm storm, open-handed politico-scientist, Nicolas L. for his pragmatic ways with intractable issues, Weiya C. for his juggling between simplicity, care for others and efficiency, and Patrick B. for the grumpy yet adorable way he fathers them all. Finally, a thousands thanks to Nédé, who'll forever have my gratitude for helping me to consciousness every morning with her home-brewed coffee.

My time at ETIS also had its lot of privileged moments with some of its inhabitants, most of them part of the Ph.D students community that hovers there. Fouad S. and his deep curiosity and knowledge over blackhat's lore, Gaël R. half teddy bear half cold genius or Romain T. and the shared tears after his own defense (dude that was moving). Finally, while he had to leave me for the greater good after a few month of collaborative work, I'd like to thank Patrick D. for introducing me with what would become the first steps of my Ph.D's stair.

How could I ever forget (not that I'd ever try) my fellow engineers from Astrium Services. Franck C, that I'd have loved even if he didn't came from the same stock than I did, Wannes and his discrete way to befriend his surroundings, Olivier K. with his mystical "I'm coming, go ahead" (to this very day I have no ideas of his whereabouts back then). My most humble thanks to my elders there: Marc C, Florence D, Claude P, Herve F, Wilfried L, Michel H, each advising me in their own way, taking time to share ideas and brainstorm on whatever random subject popped up on my desk that day.

Last but not least, I would like to thank those I chanced to work with during my Ph.D outside of any institution, amongst which Stephane A. who so willingly and throughly advised me on the operational needs regarding search and rescue. A special pensé to Dalai F, whom I first met during the Blender Conference in 2012 and with which I have had the genuine pleasure to work with since. Greetings also to all these brilliant mind I met during the various conferences of these last three years, your reading this manuscript may well alleviate your opinion of me (I'll let you decide in which way though). Thanks to all the good will of the Open Source community around the world who's daily work make life easier for hasty post-graduates in need for plug-and-pray solutions, codes, piece of softwares, etc.

Finally, I am most grateful to my family, mum and dad who educated me, my sisters and brother who saw that I remained so. Jean, I'd like you to read these lines before you leave, as I wish Geo did. In so many ways this work belongs to both of you.

Nadia, useless to tell you the role you played in the writing of this manuscript or during the eight years of schooling that brought me here. Again, so many things, so few words. Thank you for entering my life.

David

Glossary

ASon	Acoustic Sonification
AWGN	Average White Gaussian Noise
BRTF	Bone Related Transfer Function
BTS	Base Transceiver Station
CAVE	Cave Augmented Virtual Environment
CDMA	Code Division Multiple Access
CRB	Cramer Rao Bound
DCS	Digital Communication System
DF	Direction Finder
DOA	Direction Of Arrival
DoF	Degree of Freedom
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Technique
FDMA	Frequency Division Multiple Access
FIM	Fisher Information Matrix
GMSK	Gaussian Minimum Shift Keying
GPS	Global Positioning System
GRC	GNU Radio Companion
GSM	Global System for Mobile communications (also referred to as 2G)
GUI	Graphic User Interface
HCI	Human Computer Interface
HRTF	Head Related Transfer Function
ILD	Interaural Level Difference
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IQML	Iterative Quadratic Maximum Likelihood
IR	Impulse Response
ITD	Interaural Time Difference
LBS	Localization Based Services
LTE	Long Term Evolution (also referred to as 4G)
MBSon	Model Based Sonification
ML	Maximum Likelihood
MSE	Mean Squared Error
MUSIC	MUltiple SIgnal Classification
OSC	Open Sound Control
PMSon	Parameter Mapping Sonification
PP	Physical Prototype
RAM	Random Access Memory

RF	Radio Frequency
RMS	Root Mean Square
ROM	Read-Only Memory
RSS	Received Signal Strength
SDIS	Service Départemental d'Incendie et de Secours
SDR	Software Defined Radio
SIM	Subscriber Identification Module
SNR	Signal to Noise Ratio
SONAR	SOund NAvigation and Ranging
SPL	Sound Pressure Level
TDMA	Time Division Multiple Access
UCA	Uniform Circular Array
UDP	User Datagram Protocol
UHD	Universal Hardware Driver (USRP driver for Matlab)
ULA	Uniform Linear Array
ULA	Uniform Linear Array
UMTS	Universal Mobile Telecommunications System (also referred to as 3G)
USRP	Universal Software Radio Peripheral
VE	Virtual Environment
VP	Virtual Prototype
WIP	Walk-In-Place
WSF	Weighted Subspace Fitting

Contents

Chapter 1

Introduction

Contents

Avalanche, forest fire, collapsed building, shipwreck, hiking or simply a rescue operation that went wrong; there are many situations were finding someone's position suddenly becomes a critical issue. Much progress has been made in the area since the 1960s, when the first satellite constellation "Transit" was launched, enabling worldwide positioning or when John Lawton introduced the very first avalanche transceiver "Skadi". As yet, there is nothing like an optimal solution fit for every solution. To this day, the domain remains subject to active investigations through both industrial and research projects [4–7].

1.1 Context

Mobile phones and cellular networks have been considered as key enablers for tracking and positioning systems since their creation in the early 1980s [8]. Features related to Localization Based Services have progressively been integrated into telecommunication standards [9] since 1996 when US emergency responders expressed the need to trace back emergency calls made from mobile devices [10]. Their potential as positioning beacons grew as they became more popular, up until early 2014 when the International Telecommunication Union estimated the number of subscriptions around 7 billion worldwide [11], equivalent to 95.5% of the world population. It grew even further as mobile phones progressively established themselves as an ubiquitous accessory: always available, always activated, which represents an invaluable "distress beacon-like" built-in behavior.

There are two different approaches for localizing a cellphone: geolocation and goniometry. Geolocation aims at the estimation of the position of an emitter on a map (latitude/longitude) based on the signals received by several antennas positioned

around said emitter. GPS (Global Positioning System) is based on this technique, using satellite to emitter round-trip-time to estimate the emitter's position on Earth. Goniometry consists in estimating a relative direction instead of a position, generally based on the signal received by an array of antennas seen as co-localized from the emitter's point of view. Referred to as Radio Direction Finding, this technique does not in itself allow one to *localize* a cellphone; as a compass does not allow to localize the position of the North Pole if its user does not decide at some point to walk towards it. The most famous application of direction finding is probably its use in radio astronomy, unexpectedly discovered by Karl Jansky in 1932 while working at Bell Labs on the reduction of static noise in transatlantic radio transmissions [12]. Using a 30-meter array of square loop antennas operating at 20.5 MHz (rotating on Ford-T wheels in a southern New Jersey potato field), Jansky noted that for a fixed position of the array his system measured a pattern that repeated itself every 23 hours and 56 minutes. Radio astronomy was born a few weeks latter as he understood that the period of this pattern matched the length of a sidereal day and traced back the origin of this emission to, or at lest towards, the Sagittarius constellation.

Regarding the application of geolocation to rescue operations, the main issue is usually that positioning accuracy depends on network density. As the deployment of a dense private network of antennas prior to search and rescue operations is hardly possible, currently available solutions often rely on public networks [13]. Aside from binding agreements and privacy issues, cellphone geolocation based on public operators' networks is irremediably limited in accuracy by the number of base stations deployed around the search area. Achieved precision generally ranges from tens to hundreds of meters between urban and rural areas [14], if indeed their is any network coverage at all. Another critical issue with classical geolocation techniques regarding search and rescue operations is their inability to provide precise estimations in indoor and multi-storey environments. As cellphones became smartphones, solutions were designed merging WiFi, GPS, accelerometers, and other newly integrated sensors' data to improve geolocation robustness and accuracy [15]. While perfectly sensible from a technical point of view, large scale rescue operations cannot rely on these methods as they at some point require hardware or software specific devices.

Systems based on goniometry are usually favored in search and rescue applications, as they make up in portability and adaptability what they lack in range and remote monitoring as compared to geolocation. Also referred to as Direction Finders (DFs), these systems a priori rely on nothing but an active emitter, providing a somewhat progressive localization accuracy as the operator moves towards the target. Already used as a standard to localize avalanche victims since the late 1990s [16], direction finding was soon adapted to cellphone-based search and rescue operations [17]. As detailed in the next chapter, the main obstacle to using cellphones as distress beacons is that un-

like avalanche transceivers, cellphones are not programmed to continuously broadcast on a specific frequency. Available solutions can overcome this issue e.g. by deploying a portable local network to access and control every mobile phone in the search area [18].

DF designs are typically based on a directional antenna that is manually steered during the search, behaving much like a geiger counter or a metal detector [19]. Based on audio or visual feedback, the user simply follows the direction that corresponds to the maximum of received signal strength. Some more advanced designs involve a preprocessing scheme that computes the Direction Of Arrival (DOA) of a signal received by an array of antennas [20], usually coupled with a compass-like visual display. The theory behind these "automatic" DF designs is discussed in Chapter 3 and a typical implementation of both automatic and "manual" DF is detailed in Chapters 10 and 11.

Attempts have been made to design DFs with a sound-based interface yet more as an optional feature than as a performance enabler. The use of audio for manual DFs generally reduces to the mapping of the received signal strength onto the pitch or the repetition period of a sound, while it is unheard of for DOA-based DFs. Reading through previous research on *auditory displays* over recent decades, it seems however that the human auditory system is quite capable in terms of localization. For a start, humans are capable of effortlessly estimating a sounds' DOA in their surroundings with an average precision below 10 square degrees (azimut×elevation) [21]. Implementations of visual-to-auditory sensory substitution demonstrated how visually impaired users could "hear" images or video streams [22,23] to gather navigation cues from their environment. With time and training, the auditory system is capable of analyzing complex soundscapes far beyond the capacities of any modern computers. To this day, nuclear submarines still rely on their "golden ears" sonar operators to identify distant ships or topographies. Ever since World War I, these trained listeners have been able to infer a ship's speed, direction, distance, size, or even type simply by listening to its sonic signature that propagates under water. Closer at hand is this habit of ringing one's phone to find it. Without any apparent effort the brain provides information like "top left, upstairs", "on the right, behind a door", "muffled, under a blanket", simply based on spatial cues, visual context, and spectral cues which relate to how the ringing sound is filtered compared to when the phone is at hands reach.

The ambition of this thesis is to work on a portable DF design where conception is focused on sound and perception from the very start. The objective is to take advantage of the capacities of the human auditory system to propose a design which is both efficient and intuitive to be used for cellphone location during search and rescue operations.

1.2 Research Overview

This research project comprised the design, conception, and evaluation of a DF based on an audio interface, applied to the search of individuals through the localization of their mobile phone. The current techniques involved in classical DF designs are introduced and discussed regarding their application to portable navigation aids. Their performance and robustness are evaluated through both theoretical tools and practical simulations. Several sonification paradigms are proposed, concerned with various levels of DF output, from raw measured signal to preprocessed DOA. A virtual prototype is introduced and used to assess the relevance of these paradigms, based on a CAVE environment coupled with a raytracing propagation model. In the meantime, a physical prototype is implemented to assess the ecological validity of the overall study and evaluated during on-site search operations.

Three topics have been particularly elaborated in the text, related to articles and proceedings published in the course of these three years of research:

- In parallel to the DF techniques review, a method to assess the impact of antenna positioning on DOA estimation accuracy is detailed for a spherical array geometry. Based on the Fisher Information Matrix, said method is used to optimize a 3-elements array configuration based on a series of geometrical considerations.
- Prior to its use in evaluating sonification paradigms, a comparative study was designed to characterize the ecological validity of the virtual prototype. The methodology implies to compare the correlation in performance variation between two versions of the same design implemented both as a virtual and a physical prototype.
- Following test sessions with the physical prototype, an investigation on the impact of the sonification on the perception of instabilities in the received signal was conducted. The benefit of a cognitive averaging based on rhythm perception is compared to a more classical temporal smoothing function coupled with a pitch-based sonification through a topology exploration task.

1.3 Thesis Organization

The work exposed in this manuscript is organized into three parts which frame the different domains addressed in this thesis.

Part I introduces the fundamentals of DOA estimation along with a set of tools to both characterize and optimize the performance of classical DF designs. A literature review on the main DOA estimation techniques is presented in Chapter 3, followed by an introduction to the Cramer Rao Bound (CRB) in Chapter 4, used together with

Monte Carlo simulations based on propagation models as a series of tools to assess the performances of DF designs. Chapter 5 presents a case-study characterization and optimization of a lightweight head-mounted antenna array based on said tools, followed by a discussion of the limitations of advanced DOA estimation techniques applied to portable applications.

Part II presents the investigations on DF sonification design. Chapter 7 introduces concepts of sound perception along with an overview of previous research in the field of sonification with a focus on interactive Parameter Mapping Sonification (PMSon). Various sonification paradigms applied to potential DF outputs are exposed and discussed in Chapter 8. Chapter 9 presents an investigation concerning the reduction of noise perception in real-time PMSon of a noisy data streams.

Part III exposes the implementation and use of DF prototypes to assess the efficiency and general appreciation of the implemented sonification designs. Chapters 10 and 11 respectively present work related to the virtual and the physical prototype. Chapter 12 exposes the achievements related to the evaluation of the ecological validity of the virtual prototype compared to the physical prototype.

Finally, a review of the thesis results and contributions is presented in the conclusion, Chapter 14. Prior to Part I, the next chapter press,t the background context of the project. This includes a short description of the GSM protocol, focusing on its radio interface to understand how to turn a cellphone into a distress beacon.

Chapter 2

From Standard Cellphone to Rescue Beacon

Contents

As of May 2014, there were nearly 7 billion mobile subscriptions worldwide according to an estimation by the International Telecommunication Union [11].As nearly all of these cellphones are compatible with the GSM, this protocol is currently the best candidate to serve as a universal standard for rescue operations. This chapter is a brief introduction to the GSM protocol, presenting features specific to the targeted search and rescue application. The objective is to understand the different steps required to turn a standard cellphone into a controlled Radio Frequency (RF) beacon, used to guide rescuers during the search. As for any RF beacon based rescue application, the process involves (1) specifically triggering cellphones to emit on (2) known and isolated RF channels. For more information on the GSM radio interface or any other aspect of the GSM protocol, refer to the 3GPP specifications [24] or the exhaustive reviews in [25, 26].

2.1 Overview of the GSM Protocol

GSM, short for Global System for Mobile communications, is a radio telecommunication standard created in the 1980's. The term originally referred to the "Groupe Special Mobile", a European organization created in 1982 to lay the foundations of what would become the first worldwide acknowledged protocol for mobile voice-over-radio transmissions.

Table 2.1: E-GSM-900 and DCS-1800 Uplink and Downlink frequency bands.

Frequency band	Uplink channels (MHz)	Downlink channels (MHz)
E-GSM-900	880.0 - 915.0	925.0 - 960.0
DCS-1800	1 710.2 - 1 784.8	1 805.2 - 1 879.8

GSM is based on a network of BTS (Base Transceiver Station), mast-mounted antennas that relay information between subscribers handset (i.e. cellphone) and the operators core network. To register on a specific network, both handset and subscriber must be acknowledged by the operator. The subscriber is identified via an IMSI (International Mobile Subscriber Identity), a number stored on the SIM (Subscriber Identification Module), a chip card provided by the operator to its subscribers and inserted into the handset. The handset is identified via its IMEI (International Mobile Equipment Identity), a number hardcoded in the handset's motherboard.

The average BTS radio coverage area is called a cell. Every handset in a cell is said to be *attached* to its BTS. Even when the subscriber doesn't request a call, his handset is constantly exchanging messages with the attached BTS, e.g. to measure radio propagation quality. When moving from one cell to another, the handset will automatically attach itself to the BTS that presents the best reception quality, provided the BTS accepts the handset as belonging to its core network. To avoid inter-BTS radio interferences, neighboring cells are configured to operate on different frequency bands.

2.2 Focus on the Radio Interface

In Europe, GSM operates mainly on two frequency bands: E-GSM-900 and DCS-1800. It is considered to operate in *duplex mode* as the uploads (from handset to BTS) and downloads are handled through separated bands, resumed in Table 2.2.

2.2.1 Multiplexing Schemes

GSM and DCS (Digital Communication System) bands are split in 200 kHz wide frequency channels, shared amongst operators. Each RF transmission from call to traffic control[1] happens on one of these channels. The dynamic attribution of RF bands to the different transmission channels is handled though a multiplexing scheme referred to as Frequency Division Multiple Access (FDMA). The use of FDMA in GSM simply allows one to avoid situations where all the handsets of a given area would try to reach the nearest BTS on random frequencies, most likely interfering with one another in the process. To provide users with an average call quality, operators usually combine FDMA with *frequency hopping* where the network forces handsets to continuously

[1] Traffic control designates data exchanges related to network operation (e.g. call request).

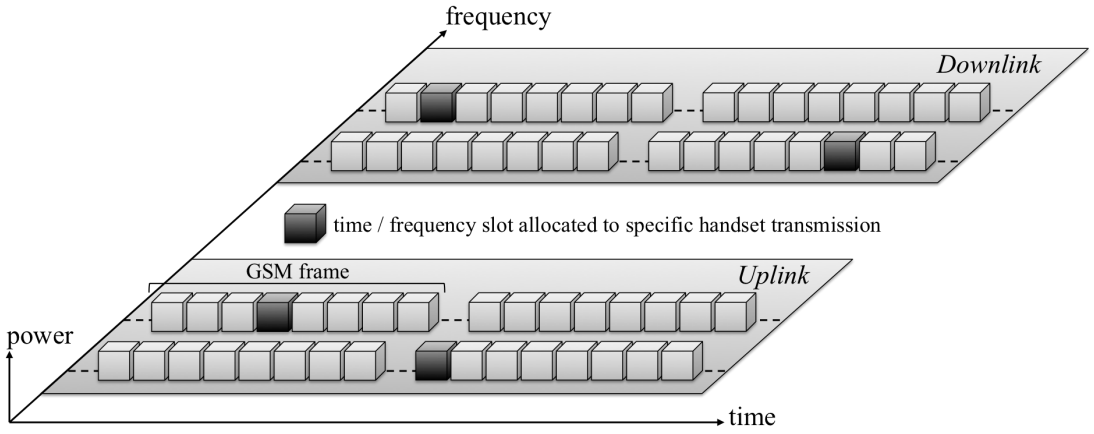


Figure 2.1: Illustration of the TDMA, FDMA and frequency hopping multiplexing schemes for uplink and downlink channels of a GSM transmission. Each cube represents a time slot on a frequency channel. A group of eight time slots is referred to as a GSM *frame*.

shift between frequency channels. Specific to voice communications, this process has been designed to improve the robustness of RF transmissions relative to narrowband interference.

The second multiplexing occurs in time, through Time Division Multiple Access (TDMA). As the average call does not require a full 200 kHz bandwidth, each frequency channel is divided into *time slots*. Each handset can access its allocated RF channel once every 8 time slots, equivalent to approximately 577 µs of transmission every 4.6 ms. Figure 2.1 illustrates both TDMA and FDMA schemes along with frequency hopping during a GSM transmission.

The last multiplexing scheme proposed in the GSM protocol is the Code Division Multiple Access (CDMA), based on data coding through pseudo random binary sequences [27]. CDMA is only applied in the US standard and in the next generation of telecommunication protocols (UMTS, LTE, etc.).

2.2.2 Modulation

The modulation specified in the GSM protocol is the Gaussian Minimum Shift Keying (GMSK). This modulation is the result of a MSK modulation [28], where binary data modulate the phase of the carrier frequency, filtered by a Gaussian filter. Rather than an amplitude modulation, the GMSK was chosen because of its robustness regarding the propagation channel and its impact on the signals' envelope. Compared to a standard MSK, the GMSK has the advantage of reducing sideband power, which in turn reduces out-of-band interference between signal carriers in adjacent frequency channels. The resulting envelope of a GMSK modulation and a typical frequency spectrum of a GSM

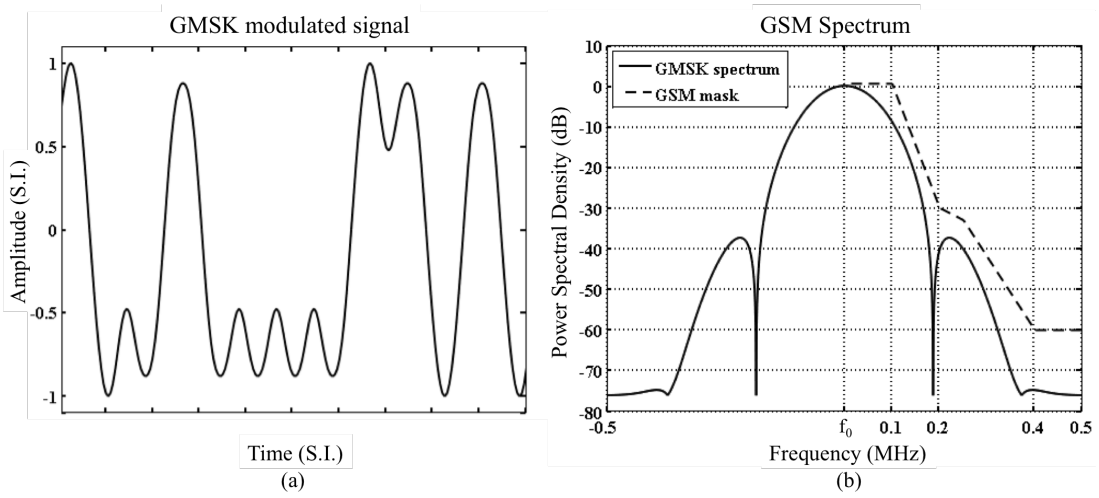


Figure 2.2: Illustration of the GMSK modulation. (a) Signal envelope after GMSK modulation on binary data. (b) GSM spectrum and GSM mask requirement based on [29], for f_0 in E-GSM-900 or DCS-1800 frequency bands of Table 2.2.

transmission are illustrated in Figure 2.2.

2.2.3 Handset States

Once switched-on, a GSM handset has basically 2 nominal states, *idle* and *dedicated*. The dedicated mode is triggered for incoming or outgoing calls, it corresponds to a handset emitting on a specific frequency channel and time slot at a given time. When in idle mode, the handset emission is reduced to a minimum. The handset in idle mode will monitor neighboring cells signal strength for potential cell re-selection, periodically check synchronization in time and frequency with the network, and stand ready to answer incoming calls. The passage from idle mode to dedicated mode is illustrated in Figure 2.3, triggered by a *paging request* sent by the BTS signaling that someone tries to reach a MS located in its cell.

2.3 Using Cellphones as RF Beacons

Discussed in the introduction of this chapter, the requirements to turn a handset into a RF beacon become (1) to be able to specifically trigger the handset into dedicated mode (2) on a known time slot and frequency channel. An additional step is eventually required to gather information on the handsets present in the search area.

The two known solutions to reach these requirements are to control the handset from its home network or to hack the link between the handset and its operator (*man in the middle* or *ghosting* attack, see [30]). While obviously not discussed hereafter, the second solution is an option that does not require much more than an average

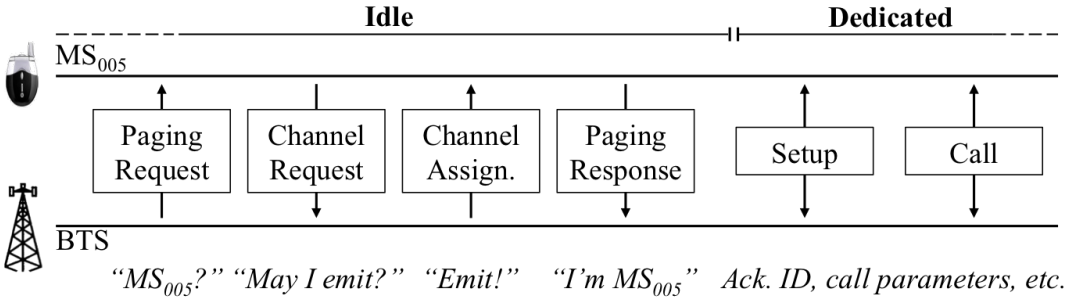


Figure 2.3: Transition from idle to dedicated mode triggered by a paging request related to an incoming call (oversimplified incoming call flow). The term dedicated refers to the nature of the channel used for the transmission rather than the RF activity of the MS.

laptop wired to an RF front end [31] as far as the GSM protocol is concerned. Next generation telecommunication protocols like 3G-UMTS or 4G-LTE implement far more robust protection algorithm in terms of cryptography or identification procedures to prevent such attacks [32].

The "control of the handset as its operator" is the only sensible and long-term viable solution for the considered application. Once one controls the handset network, its enough to *page* it (e.g. initiate an incoming call) based on its IMEI or the IMSI of its SIM card to force the handset into dedicated mode onto a given time slot and frequency channel. Using a tuned receiver listening to the same frequency channel and time slot, it is then possible to receive the RF signal of this specific handset and estimate its Direction Of Arrival relative to the receiver. Such control over every handset in the search area is assumed in the application context of the thesis as further details on this procedure are beyond the scope of this manuscript.

A typical method to switch every GSM handset of the search area onto a private controlled network is illustrated in Figure 2.4. Similar to the solution detailed in [19], this jammer-based implementation relies on a procedure already hardcoded as a standard into every GSM handset that privileges connections to any available network in the absence of any identified home network. **Note that this process is not legal, whatever the country, unless approved by a competent authority**.

2.4 Conclusion

This chapter briefly described the fundamentals of the GSM protocol, focusing on aspects related to the search and rescue application considered in this manuscript. It has been shown that to turn a GSM handset into a RF beacon, one needs to be able to force

it into a dedicated mode onto a known and isolated time slot and frequency channel. The most sensible and viable solution, assumed in the reminder of this manuscript, is to control the GSM network over the search area.

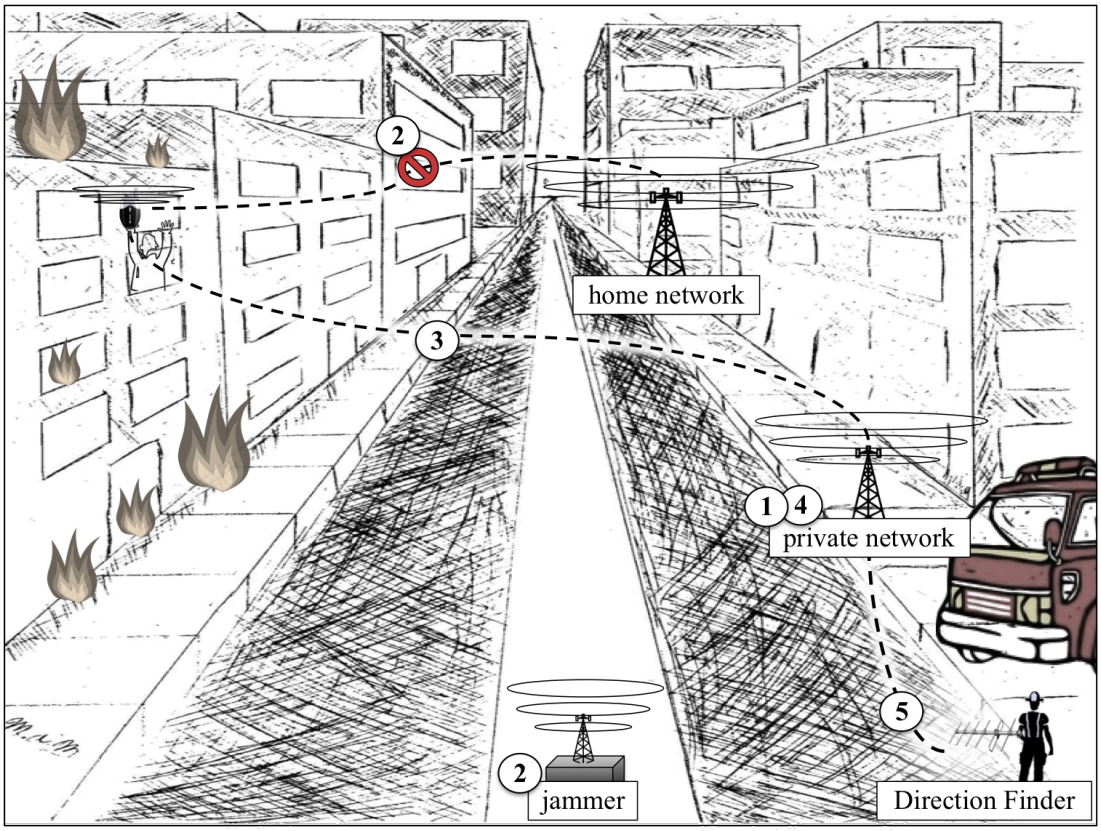


Figure 2.4: Illustration of a jammer-based scenario to turn a GSM handset into a RF beacon for DF aided search. (1) Deployment of a private GSM network on frequency bands not used by official operators in the search area. (2) Jamming of every other GSM frequency band to deny nearby handsets access to their home networks. (3) Grant access to the private network for all the handsets in the search area. GSM handset are programmed to request network attachment to any available network if they cannot reach their home network. (4) Get a list of handsets attached to the private network, iteratively switch them in dedicated mode on known frequency channel and time slot. (5) Meanwhile, iteratively forward the corresponding radio parameters to tune the DF and localize the handsets.

Part I

Direction Of Arrival Estimation

The work presented in this part of the manuscript concerns research related to Direction Of Arrival estimation. An overview of the classical techniques is provided followed by an introduction to a set of tools used to assess the performance of a DOA estimator. A typical optimization process is detailed in the last chapter, focusing on a lightweight antenna array designed for search and rescue operations. The term "Direction Finder" is deliberately replaced by the notion of DOA estimator in the following chapters to distinguish between the combination of a DOA estimation technique and an antenna array (DOA estimator), and the combination of a DOA estimator with a user interface (Direction Finder) discussed in Parts II and III.

Chapter 3

Introduction to Direction Of Arrival Estimation

Contents

Radiogoniometry, i.e. the Direction Of Arrival (DOA) estimation of received radio waves is nearly as old a science as radio transmission itself. Its application ranges from spectrum sensing to civilian rescue operations [33, 34]. The term DOA estimation generally refers to the application of a DOA estimation technique on the output of an array of sensors. Countless techniques have been designed during the last decades, gradually improving accuracy, efficiency, and robustness of the estimations [35]. This chapter does not intend to provide an exhaustive description of each technique, but rather to focus the reader on DOA estimation fundamentals. In a nutshell: how to estimate a DOA, what are the main techniques, what are the main issues, and finally, what are their potential solutions. For more detailed reviews of DOA estimation techniques and related applications, refers to [35–37].

3.1 Introduction

The combination of a DOA estimation technique and a array of sensors is referred to as a *DOA estimator*. The characterization of a DOA estimator generally focuses on the accuracy of its estimations, its ability to segregate sources impinging from close DOAs (i.e. its *resolution*), its robustness to noisy or correlated signals, and its computational complexity. Unless otherwise specified, the arrays of sensors considered in this chapter

will always refer to geometries composed of several omnidirectional antennas. Note that DOA estimation is not limited to Radio Frequency (RF) nor electromagnetic propagation: the overall theory discussed hereafter is perfectly compliant with e.g. sound waves propagation [38].

Nearly every DOA estimation technique is based on the observation of a shift in time, or equivalently in phase, between signals received by the different antennas of an array[1]. A simple analogy is to consider leaves (antennas) floating at the surface of a pond (RF propagation medium) in which a stone is thrown (RF source to be localized). By observing the motion patterns of the leaves, one can infer the rough direction from which came the ripples, as the delay between the respective oscillations of the leaves depends on said direction. For a computer, this process is equivalent to the simple correlation-based DOA estimation discussed in the following section. The need for more complex techniques eventually arises if we consider that the ripples disturbing the leaves can also come from the wind (noise in the propagation channel), reflections off the shore (multi-path propagation), and other stones simultaneously thrown in the pond (multi-sources propagation). The analogy also suggests that the respective positions of the leaves/antennas compared to one another will impact the performance of the DOA estimation.

DOA estimation techniques, and more generally *power spectrum* estimation techniques, can be broadly classified into two categories: parametric and non-parametric methods [39]. Parametric methods can also be classified as *model-based*, i.e. they assume the data to be issued from a known model with unknown parameters (e.g. the DOA). Non-parametric methods, also referred to as *spectral-based*, do not assume any model for the measured dataset, usually relying on Fourier-like spectrum analysis. It goes without saying that, supposing a realistic enough modelization, the performance of parametric methods eventually exceed non-parametric's.

The following sections present the generic signal model used in this chapter along with three of the most common classes of DOA estimation techniques: Beamforming, Subspace-based decomposition, and Maximum Likelihood (ML). Beamforming is a class of spectral-based techniques, subspace decomposition and ML are typical model-based techniques.

3.2 Signal Model and Antenna Array

The following scenario is assumed and maintained throughout the following sections:

[1] The inter-antennas level difference is seldom used in RF DOA estimation as the decrease in received signal strength between the elements of the antenna array is most of the time below the measurement threshold, c.f. "Far-field assumption" in the next section.

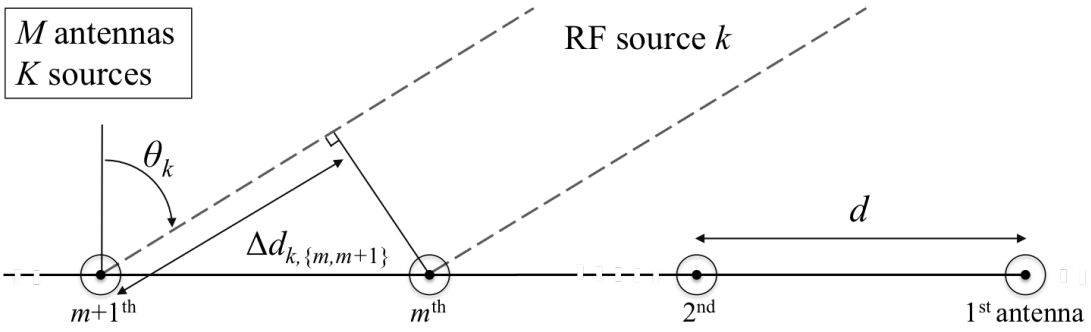


Figure 3.1: M-element ULA of step size d, impinged by K sources from DOAs $\{\theta_1, ..., \theta_k\}$. $\Delta d_{k,\{m,m+1\}}$ represents the difference of path length between array elements m and $m+1$ for impinging source k.

Isotropic and linear transmission medium: The transmission medium between the RF sources and the antenna array is assumed to be isotropic and linear. As such, the propagation property of the wave is not impacted by its DOA. RF signals traveling through the medium and received by the antennas can be modeled as a linear wavefronts superposition.

Far-field assumption: or plane wave assumption. The DOA of the received signals does not depend on the antenna position in the array. The RF source is assumed to be distant enough from the array for its wavefront to be modeled as parallel planes rather than concentric circles. A rule of thumb is that the far-field assumption is true for distances larger than $2\mathrm{D}^2/\lambda$, with D the typical dimension of the array and λ the wavelength of the RF source. D is usually approximated as the distance between two antennas in the array. For a typical DOA estimator, $\mathrm{D} = \lambda/2$ and $\lambda \leqslant 34$ cm for GSM signals which validates the far-field assumption for RF source-to-array distances greater than 15 cm.

Narrowband and uncorrelated sources: All sources have their frequency content concentrated in the vicinity of a single carrier frequency f (also referred to as *monochromatic* sources). Signals impinging on the array are considered uncorrelated. This last hypothesis implies that the propagation channel does not generate multi-path propagation from RF sources to array. As discussed in Chapter 4, the uncorrelated sources assumption does not hold for most propagation channels.

AWGN channel: An Average White Gaussian Noise (AWGN) of variance σ_n^2 is added to the received signals to model the impact of the propagation channel.

To illustrate the main DOA techniques, let us assume the case study illustrated in Figure 3.1, where K uncorrelated narrowband RF sources impinge on a Uniform Linear

Array (ULA) composed of M antennas, with $M > K$, from DOAs $\{\theta_1, ..., \theta_k\}$. A ULA is an array with its antennas evenly distributed on a line, where the distance between two neighboring antennas is defined as the ULA step size, noted d in Figure 3.1. Sources and antenna array are assumed coplanar to constrain the study to a 1D DOA estimation, antennas are assumed as omnidirectional with unity gain. The assumed ULA array and the coplanar hypothesis allow for a simplification of mathematical formulations, the techniques presented in the next sections perfectly extend however to random 3D array geometry for 2D DOA estimation (i.e. estimation of both azimuth and elevation).

Note that correlated or non-monochromatic sources would prevent some of the assumptions made in the remainder of this section. The impact of multi-path propagation on classical estimation techniques is addressed in the case-study of Chapter 5.

As seen in Figure 3.1, the signal received at array $m+1^{\text{th}}$ element is a time shifted version of the one received at the m^{th} element, itself shifted compared to the $m-1^{\text{th}}$, etc. until the 1^{st} element. Assuming the signal received at antenna m from source k is noted $s^{\mathbb{R}}_{k,m}(t)$, where $\mathbb{R}$ indicates that s is real valued, and simplifying the notation to $s^{\mathbb{R}}_{k}(t)$ for the 1^{st} element, the signal received at the m^{th} element can be expressed as

$$s^{\mathbb{R}}_{k,m}(t) = s^{\mathbb{R}}_{k}(t - \Delta t_{k,m}) \tag{3.1}$$

where $\Delta t_{k,m}$ represents the time delay between 1^{st} and m^{th} elements, with

$$\Delta t_{k,m} = (m-1)\frac{\Delta d_{k,\{m,m+1\}}}{c} = (m-1)\frac{d\sin\theta_k}{c} \tag{3.2}$$

$\Delta d_{k,\{m,m+1\}}$ being the difference of path length between array elements m and $m+1$, θ_k the DOA of source k and c the speed of light in the propagation medium. Note that the selection of a specific element as a time reference simplifies result formulation yet does not impact on their generalizability.

Equation (3.1) is usually manipulated in its complex form, were phase shifts can easily be factorized. In the remaining of this section, $s_k(t) = s^{\mathbb{C}}_k(t)$ refers to the complex envelope of the k^{th} signal of the form $s_k(t) = \alpha_k(t)e^{j\beta_k(t)}$. From (3.1) and (3.2), the complex form of signal k received at antenna element m is

$$s_{k,m}(t) = s_k(t)e^{-j(m-1)(2\pi f/c)d\sin\theta_k} = s_k(t)e^{j(m-1)\mu_k}$$

where f is the carrier frequency, common to all signals and μ_k, often referred to as the signal *spatial frequency*, is a function of θ_k defined for the considered ULA as

$$\mu_k = \frac{-2\pi f}{c}d\sin\theta_k = \frac{-2\pi}{\lambda}d\sin\theta_k$$

where $\lambda = c/f$ designs the common carrier wavelength. The instantaneous array output vector, noted $\mathbf{x}(t)$, can then be expressed as the sum of the K incident signals on the M antennas, each corrupted by channel noise $n_m(t)$

$$\begin{aligned}\mathbf{x}(t) &= \sum_{k=1}^{K}\left(\sum_{m=1}^{M} s_k(t)e^{j(m-1)\mu_k} + n_m(t)\right)\\ &= \sum_{k=1}^{K} s_k(t)\mathbf{a}(\mu_k) + \mathbf{n}(t)\\ &= \mathbf{A}(\mu_1, ..., \mu_k)\mathbf{s}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s} + \mathbf{n}\end{aligned} \tag{3.3}$$

where $\mathbf{s}(t) = [s_1(t), s_2(t), ..., s_K(t)]^T$ and $\mathbf{n}(t) = [n_1(t), n_2(t), ..., n_M(t)]^T$ are respectively the instantaneous signal and noise vectors. The $\mathbf{a}(\mu_k)$ or equivalently the $\mathbf{a}(\theta_k)$ for $k \in [\![1, K]\!]$ are referred to as the array steering vectors. Together they form the columns of the $M \times K$ steering matrix $\mathbf{A}$

$$\mathbf{a}(\mu_k) = \begin{pmatrix} 1 \\ e^{j\mu_k} \\ \vdots \\ e^{j(M-1)\mu_k} \end{pmatrix} \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{j\mu_1} & e^{j\mu_2} & \dots & e^{j\mu_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\mu_1} & e^{j(M-1)\mu_2} & \dots & e^{j(M-1)\mu_K} \end{pmatrix} \tag{3.4}$$

As the M time-shifted versions of the K signals are correlated compared to the propagation related AWGN, most of the DOA estimation techniques will introduce the notion of *spacial covariance matrix*: the matrix form of the standard cross-correlation often used to detect time delay in a noise-corrupted sequence. The spatial covariance matrix $\mathbf{R}_x$ of input signal $\mathbf{x}(t)$ is defined as

$$\mathbf{R}_x = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} \tag{3.5}$$

where $E\{\cdot\}$ denotes the statistical expectation and $\mathbf{x}^H(t)$ the Hermitian conjugate (conjugate transpose) of the vector $\mathbf{x}(t)$. $\mathbf{R}_x$ is replaced by its estimate $\hat{\mathbf{R}}_x$ in practical applications, as the time constraint limits the number of samples available per estimation:

$$\mathbf{R}_x \approx \hat{\mathbf{R}}_x = \sum_{n=1}^{N} \mathbf{x}[t_n]\mathbf{x}^H[t_n] \tag{3.6}$$

Actually, correlation is in itself a non-parametric DOA estimation technique. The Cauchy-Schwarz inequality states that $\mathbf{a}^H(\theta)\mathbf{a}(\theta_k)$ as defined in (3.4) has a maximum for $\theta = \theta_k$. For simple scenarios, assuming uncorrelated and narrow band sources, the function

$$P_{corr}(\theta) = \mathbf{a}^H(\theta)\mathbf{x}(t) \tag{3.7}$$

will present K main peaks positioned at sources DOAs $\{\theta_1, ..., \theta_K\}$. $P_{corr}(\theta)$, or more generally $P_{technique}(\theta)$ in the following sections is referred to as the *spatial power spectrum* of the received signal. Note that, due to the linear equi-spaced geometry of the ULA, the elements in $\mathbf{a}^H(\theta)$ can be interpreted as Fourier coefficients and (3.7) as a discrete Fourier transform of the data vector $\mathbf{x}(t)$, which K main coefficients are the spatial frequencies of $P_{corr}(\theta)$.

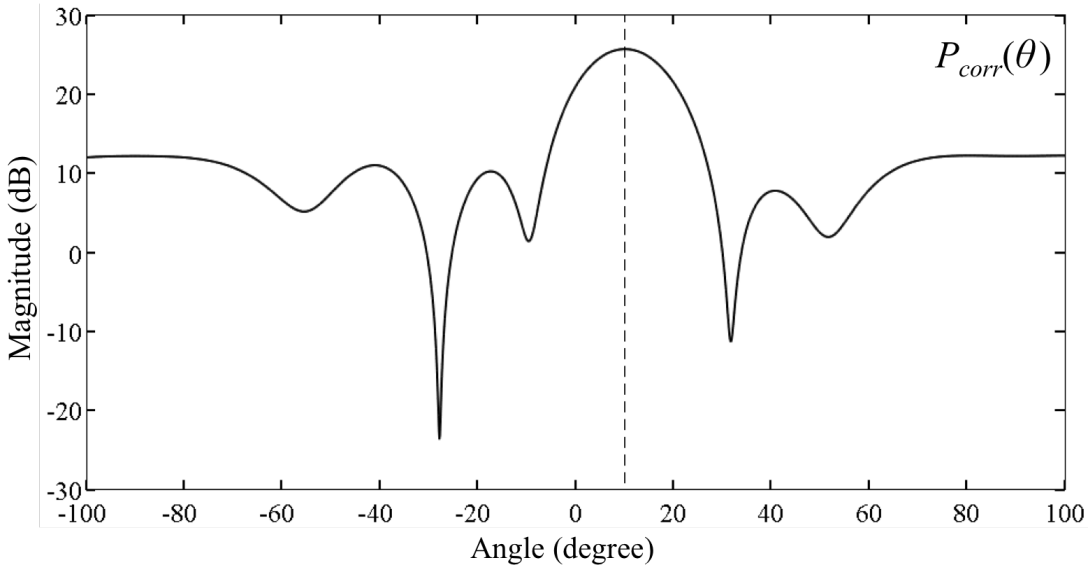


Figure 3.2: Plot of the spatial power spectrum $P_{corr}(\theta)$ associated to the correlation technique with a single source of DOA $\{\theta = 10^\circ\}$. The estimation is correct yet the low resolution of the correlation technique limits its application in realistic scenarios.

To illustrate the different resolutions of DOA estimation techniques throughout the following sections, comparative plots of their spatial power spectrums will be presented, all based on a similar simulation scenario. This scenario assumes a $M = 6$ elements ULA of step $d = \lambda/2$ and omnidirectional gains $= 1$, impinged by K sources from DOAs $\{\theta_1, ..., \theta_k\}$. Source and noise signals are both simulated through AWGN (but the same source hits every antenna while the noise is antenna specific so as to remain uncorrelated). The Signal to Noise Ratio (SNR) of each source is fixed to 1 dB, the DOA is estimated based on a 100 sample wide array output vector $\mathbf{x}(t)$. As an example, $P_{corr}(\theta)$ spectrum is plotted in Figure 3.2 for a single source with DOA $\{\theta = 10^\circ\}$.

3.3 Beamforming Techniques

The basic principle of beamforming techniques is to *steer* the antenna array in one direction at a time and measure the total Received Signal Strength (RSS), which reaches maximums for steering angles that coincide with the signals' DOAs. The steering of the array towards a given direction is achieved by introducing specific delays on the received signal of every array element through a complex weighting vector $\mathbf{w}(\theta)$, much like the correlation method. These delays are designed to combine the antenna outputs so that the signals from a given direction are added coherently. The principle may be easier to understand for an emitting array, where the same delay configuration enables one to emit *in-phase* signals to form constructive interferences between antennas' signals in a given direction.

Initially based on a fixed delay hardcoded in the array design[2] coupled with mechanical steering, modern beamforming arrays are steered electronically. The array output signal $y(t, \theta)$ is a time-weighted sum of the antennas outputs, expressed with $\mathbf{w}(\theta)$ as

$$y(t, \theta) = \mathbf{w}^H(\theta)\mathbf{x}(t)$$

The aim of Beamforming-based DOA estimation is the determination of the steering angle(s) for which this output signal's power, or equivalently the beamforming spatial power spectrum $P_{Beamforming}$, reaches a local maximum. Assuming a signal's power estimation based on an average over a finite number of snapshots N, $P_{Beamforming}$ can be expressed as

$$\begin{aligned} P_{Beamforming}(\mathbf{w}) &= \frac{1}{N}\sum_{n=1}^{N} |y[t_n]|^2 = \frac{1}{N}\sum_{n=1}^{N} \mathbf{w}^H\mathbf{x}[t_n]\,\mathbf{x}^H[t_n]\mathbf{w} \\ &= \mathbf{w}^H\left(\frac{1}{N}\sum_{n=1}^{N}\mathbf{x}[t_n]\mathbf{x}^H[t_n]\right)\mathbf{w} = \mathbf{w}^H\hat{\mathbf{R}}_x\mathbf{w} \end{aligned} \tag{3.8}$$

where $\hat{\mathbf{R}}_x$ is the estimate of the spatial covariance matrix of $\mathbf{x}[t_n]$ over N samples, as defined in (3.6). The notation $\mathbf{w} = \mathbf{w}(\theta)$ is adopted for the sake of readability.

The most straightforward beamforming technique is the *Bartlett* beamformer [40], often referred to as the "conventional beamformer". Its weight vector is defined as

$$\mathbf{w}_{Bartlett}(\theta) = \frac{\mathbf{a}(\theta)}{\sqrt{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}} \tag{3.9}$$

[2] e.g. adjusting the delay through wires length

for steering vector $\mathbf{a}(\theta)$ defined in (3.4). Note that the denominator in (3.9) is simply a normalization factor, equal to $\sqrt{M}$ for a M-element ULA array. From (3.9) and (3.8) it follows that

$$P_{Bartlett}(\theta) = \frac{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_x\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}$$

The *Capon* beamformer [41] is often preferred to the conventional beam former as it has a better DOA estimation resolution, i.e. a better capacity to segregate two RF sources impinging from nearby DOAs. Capon's generalized beamformer algorithm improves Bartlett beamformer through the formation of nulls in every direction but the one observed, in order to minimize the contribution of signals impinging from these DOAs to the spatial power spectrum (the number of available nulls is naturally related to the number of antennas M). Its weight vector is defined by

$$\mathbf{w}_{Capon} = \frac{\hat{\mathbf{R}}_x^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_x^{-1}\mathbf{a}(\theta)} \tag{3.10}$$

defined by minimizing the array output power while using the remaining degrees of freedom to constrain the gain in the steering direction to be unity, i.e.

$$\min_{\mathbf{w}}\big[P_{Beamforming}(\mathbf{w})\big] \text{ subject to } \mathbf{w}^H(\theta)\mathbf{a}(\theta) = 1$$

for example via the Lagrange multiplier method (see derivation in [42]). Substituting (3.10) in (3.8) gives $P_{Capon}(\theta)$, the spatial power spectrum of the Capon beamformer [3]

$$P_{Capon}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\hat{\mathbf{R}}_x^{-1}\mathbf{a}(\theta)} \tag{3.11}$$

To compare the resolution of both beamforming techniques, the simulation scenario introduced in Section 3.2 is applied to the DOA estimation of 3 uncorrelated sources of respective DOAs $\{\theta_1, \theta_2, \theta_3\} = \{-40°, 10°, 25°\}$. Resulting spatial power spectrums are reported in Figures 3.3 and 3.4.

Beamforming techniques are generally seen as having light computational complexity while being not very robust to the presence of noise and interferences in the received signal. Improved algorithms such as the Capon's beamformer will often involve matrix inversion-like operations, increasing the estimation resolution and robustness to noise as well as its computational cost.

[3] The derivation of (3.11) from (3.8) and (3.10) uses the relation $(\mathbf{R}_x^{-1})^H = \mathbf{R}_x^{-1}$, as $\mathbf{R}_x$ is an invertible hermitian matrix by construction. It can be proven that for large values of N, $\hat{\mathbf{R}}_x$ shares this property.

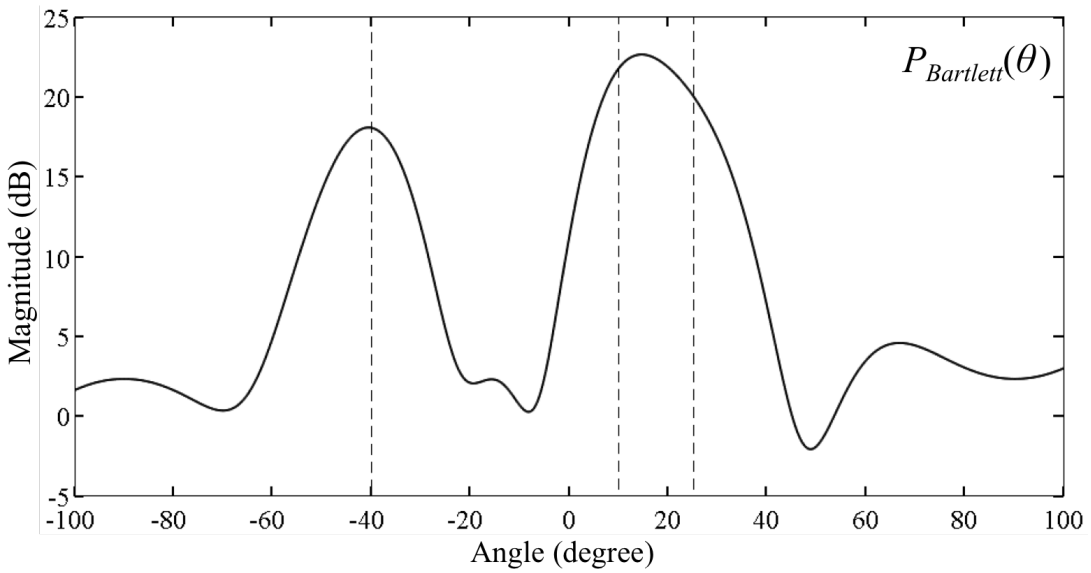


Figure 3.3: Plot of the spatial power spectrum $P_{Bartlett}(\theta)$ associated to the conventional beamformer technique for DOAs $\{\theta_1, \theta_2, \theta_3\} = \{-40^\circ, 10^\circ, 25^\circ\}$. The estimate of the left hand source DOA at -40° is correct but $P_{Bartlett}(\theta)$ resolution is not high enough to segregate sources at 10° and 25°.

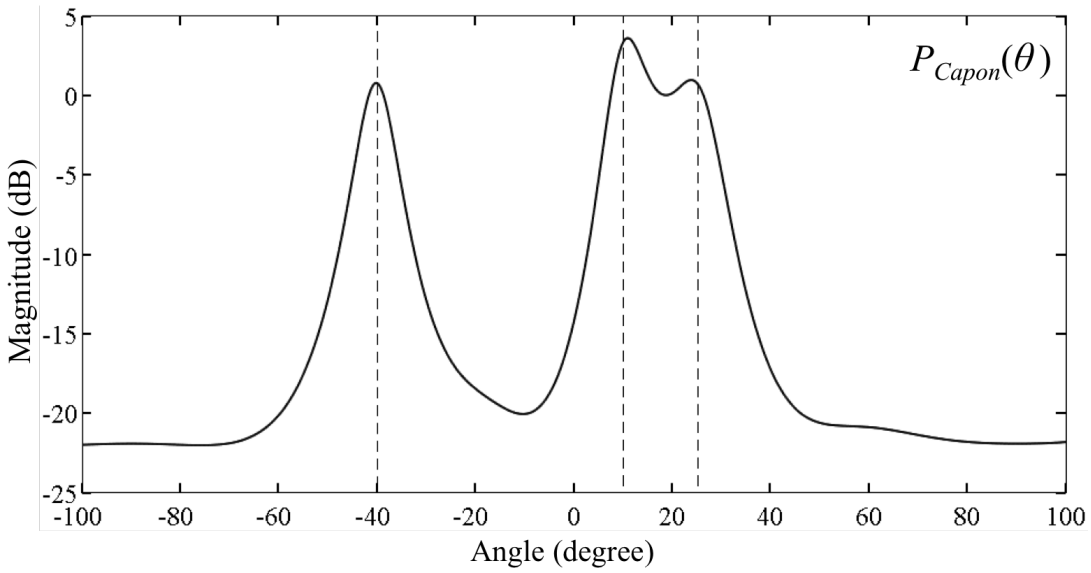


Figure 3.4: Plot of the spatial power spectrum $P_{Capon}(\theta)$ associated to the Capon beamformer technique for DOAs $\{\theta_1, \theta_2, \theta_3\} = \{-40^\circ, 10^\circ, 25^\circ\}$. It correctly segregates the three sources at -40°, 10° and 25° compared to Bartlett beamformer.

3.4 Subspace Decomposition Techniques

Research on beamforming have been progressively discarded for Subspace based techniques since their application to DOA estimation in the early 1980s [43,44]. Aside from being more efficient and offering higher resolutions, estimations based on subspace decomposition proved to be *consistent* compared to those issued from beamforming techniques. An estimator is said to be consistent if it converges to the true value of the estimated parameter when the number of measured data tends toward infinity. The derivations introduced in this section will be focused on the MUltiple SIgnal Classification proposed by R. Schmidt in 1986 [45], as the *MUSIC* technique simply and elegantly illustrates the concept of subspace-based DOA estimation.

As with most adaptive techniques[4], MUSIC involves the spatial covariance matrix $\mathbf{R}_x$ of the received signal $\mathbf{x}(t)$ defined in (3.5). MUSIC relies on the fact that $\mathbf{R}_x$ can be expressed in a base of orthogonal vectors where all but K (number of sources, assuming $K < M$ the number of antennas) belong to what is referred to as the *noise subspace*. Defining an orthogonal projection onto this subspace, the simple application of the correlation method of (3.7) with the projection of the steering vector $\mathbf{a}(\theta)$ considerably improves the resolution of the estimations.

The following derivation first addresses $\mathbf{R}_x$ diagonalization that leads to the creation of noise and signal subspaces. It then introduces the so-called MUSIC spectrum (i.e. spatial power spectrum) that takes advantage of the orthogonality between the $\mathbf{a}(\theta_k)$ steering vectors related to DOAs θ_k for $k \in [\![1, K]\!]$ and the noise subspace.

Based on (3.3) and (3.5), $\mathbf{R}_x$ can be expressed as

$$\begin{aligned}\mathbf{R}_x &= E\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\} = E\left\{\big(\mathbf{A}\mathbf{s}(t)+\mathbf{n}(t)\big)\big(\mathbf{A}\mathbf{s}(t)+\mathbf{n}(t)\big)^H\right\} \\ &= \mathbf{A}E\left\{\mathbf{s}(t)\mathbf{s}^H(t)\right\}\mathbf{A}^H + E\left\{\mathbf{n}(t)\mathbf{n}^H(t)\right\} \\ &= \mathbf{A}\mathbf{R}_\mathbf{s}\mathbf{A}^H + \sigma_n^2\mathbf{I}_M\end{aligned} \tag{3.12}$$

where $\mathbf{R}_\mathbf{s} = E\left\{\mathbf{s}(t)\mathbf{s}^H(t)\right\}$ is the signal covariance matrix, σ_n^2 the common noises variance and $\mathbf{I}_M$ the $M \times M$ identity matrix. Note here that the intermediate terms that concern the cross expectation of $\mathbf{s}(t)$ and $\mathbf{n}(t)$ are null since signal and interferences are assumed to be uncorrelated.

It can be shown that a full column rank matrix $\mathbf{A}$ ($\mathbf{A}$ is full column rank since it is a Vandermonde matrix with linearly independent rows) and a nonsingular matrix $\mathbf{R}_s$ guarantee that the matrix $\mathbf{A}\mathbf{R}_\mathbf{s}\mathbf{A}^H$ is positive semidefinite with rank K when

[4] i.e. data-dependent, adapting its coefficients, weights, etc. to the incoming data to optimize the estimation performance

there are less incident signals K than antenna M . The fact that $\mathbf{AR_sA}^H$ is positive and semidefinite guarantees that it is diagonalizable and has K non-zero and $M-K$ zero real-valued eigenvalues. Lets define the $M \times M$ change-of-basis matrix $\mathbf{Q} = [\mathbf{Q}_s\mathbf{Q}_n]$ with the matrices $\mathbf{Q}_s$ and $\mathbf{Q}_n$ respectively composed of the non-zero and zero eigenvalues-related eigenvectors. $\mathbf{Q}_n$ represents the noise subspace, $\mathbf{Q}_s$ the signal subspace.

$\mathbf{AR_sA}^H$ can then be expressed as function of a diagonal matrix $\mathbf{D}_s$

$$\mathbf{AR_sA}^H = \mathbf{QD}_s\mathbf{Q}^H$$

were $\mathbf{D}_s = \text{diag}(\lambda_1, ..., \lambda_K, 0_{K+1}, ..., 0_M)$, with $\{\lambda_1, ..., \lambda_K\} > 0$ the non-zero eigenvalues of $\mathbf{AR_sA}$. By construction, any column vector $\mathbf{q_{n}}_i$ of the noise subspace $\mathbf{Q}_n$ is orthogonal to $\mathbf{AR_sA}$ (because of its related 0 eigenvalue), i.e.

$$\mathbf{AR_sA}^H\mathbf{q_{n}}_i = 0 \tag{3.13}$$

Since $\mathbf{A}$ has a full rank and $\mathbf{R}_s$ is nonsingular (i.e. invertible), (3.13) is equivalent to

$$\mathbf{A}^H\mathbf{q_{n}}_i = 0$$

The fact that $\mathbf{A}^H\mathbf{q_{n}}_i = 0 \ \forall i \in [\![1, M-K]\!]$ indicates that the eigenvectors corresponding to the $\mathbf{R}_s$ noise subspace (of dimension $M-K$) are orthogonal to the K steering vectors $\mathbf{a}(\theta_k)$ that compose the columns of $\mathbf{A}$:

$$\{\mathbf{a}(\theta_1), ..., \mathbf{a}(\theta_K)\} \perp \left\{\mathbf{q_{n}}_1, ..., \mathbf{q_{n}}_{M-K}\right\} \tag{3.14}$$

This observation forms the cornerstone of almost all of the subspace-based methods. It means that one can estimate DOAs θ_k through a simple identification of θ values for which the steering vector $\mathbf{a}(\theta)$ is orthogonal to the noise subspace of $\mathbf{R_s}$.

In practical applications, it is not possible to compute $\mathbf{R_s}$, nor its estimate. Only an estimate of $\mathbf{R}_x$ is available, based on the received $\mathbf{x}(t)$. As it is, both $\mathbf{AR_sA}^H$ and $\mathbf{R}_x$ share the same eigenvectors: based on (3.12) and (3.13), $\forall i \in [\![1, M-K]\!]$

$$\begin{aligned}
\mathbf{R}_x\mathbf{q_{n}}_i &= (\mathbf{AR_sA}^H + \sigma_n^2\mathbf{I}_M)\mathbf{q_{n}}_i \\
&= (\mathbf{AR_sA}^H)\mathbf{q_{n}}_i + \sigma_n^2\mathbf{I}_M\mathbf{q_{n}}_i \\
&= (\mathbf{AR_sA}^H - 0.\mathbf{I}_M)\mathbf{q_{n}}_i + \sigma_n^2\mathbf{I}_M\mathbf{q_{n}}_i \\
&= 0 + \sigma_n^2\mathbf{I}_M\mathbf{q_{n}}_i
\end{aligned}$$

which can be rewritten in the canonical form

$$(\mathbf{R}_x - \sigma_n^2 \mathbf{I}_M)\mathbf{q}_{\mathbf{n}i} = 0$$

The same manipulation on signal subspace related eigenvectors gives, $\forall i \in [\![M - K, M]\!]$

$$(\mathbf{R}_x - (\lambda_k + \sigma_n^2)\mathbf{I}_M)\mathbf{q}_{\mathbf{s}i} = 0$$

$\mathbf{R}_x$ hence shares $\mathbf{R}_s$ eigenvectors and can be diagonalized as $\mathbf{D}_x = diag(\lambda_1 + \sigma^2, ..., \lambda_K + \sigma^2, \sigma^2, ..., \sigma^2) \in \mathcal{M}^{M\times M}$ using the same matrix $\mathbf{Q} = [\mathbf{Q}_s \mathbf{Q}_n]$

$$\mathbf{R}_x = \mathbf{Q}\mathbf{D}_x\mathbf{Q}^H$$

In practice, the noise subspace $\mathbf{Q}_n$ is computed from the estimate $\hat{\mathbf{R}}_x$, selecting its eigenvectors related to the $M - K$ smallest eigenvalues. Since K is usually unknown, the process requires at first the determination of the number of "smallest" eigenvalues roughly identical to one another (equal to σ^2). Once the noise subspace is created, the MUSIC spectrum

$$P_{MUSIC}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{Q}_n\mathbf{Q}_n^H\mathbf{a}(\theta)} = \frac{1}{|\mathbf{a}^H(\theta)\mathbf{Q}_n|^2} \tag{3.15}$$

is used for DOA estimation. Because of the orthogonality in (3.14), the $P_{MUSIC}(\theta)$ denominator $|\mathbf{a}^H(\theta)\mathbf{Q}_n|^2$ is null and $|\mathbf{a}^H(\theta)\hat{\mathbf{Q}}_n|^2$ tends toward zero for $\theta \in \{\theta_1, ..., \theta_K\}$.

To evaluate the efficiency of the MUSIC estimation technique, the simulation scenario introduced in Section 3.2 is applied to the DOA estimation of 4 uncorrelated sources of respective DOAs $\{\theta_1, \theta_2, \theta_3, \theta_4\} = \{-50°, -40°, 10°, 25°\}$. The resulting spatial power spectrum is reported in Figure 3.5, illustrating the high resolution of MUSIC compared to both Capon and Bartlett beamformers introduced in the previous section.

This section would not be complete without mentioning the class of estimation techniques known as ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) [46]. ESPRIT requires a property called the *shift invariance* of the array, i.e. geometries that are composed of at least two identical subarrays, a requirement not prohibitive in practical applications as many of the commonly employed array geometries exhibit these invariances.

The first asset of ESPRIT is that it does not require the array manifold steering vectors to be precisely known. Not mentioned earlier, slight differences between modeled and real array geometry (e.g. inter-antenna distance d for a ULA) for techniques such as beamforming or MUSIC have a severe impact on the accuracy of the estimations. ESPRIT has also been designed to naturally integrate a side algorithm known as forward-backward averaging, usually implemented as a pre-processing scheme to

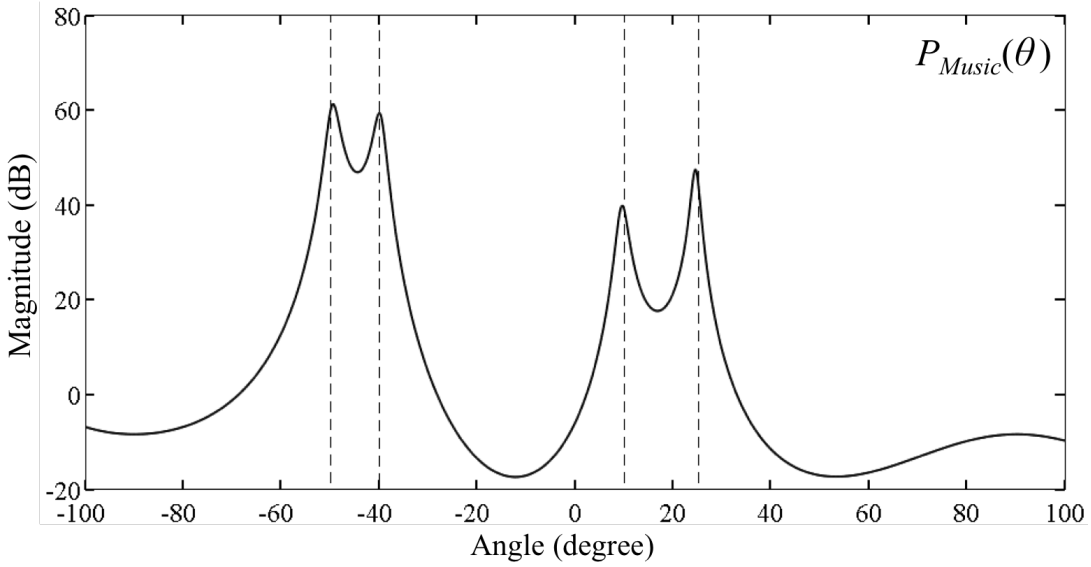


Figure 3.5: Plot of the spatial power spectrum $P_{MUSIC}(\theta)$ associated to the MUSIC estimation technique for DOAs $\{\theta_1, \theta_2, \theta_3, \theta_4\} = \{-50°, -40°, 10°, 25°\}$. It correctly segregates the four sources at $-50°$, $-40°$, $10°$, and $25°$ displaying narrow maximums compared to beamforming techniques due to the noise subspace orthogonality.

improve the robustness of a DOA estimator regarding estimation on correlated signals [37]. ESPRIT takes advantage of the shift invariance of the array to apply this forward-backward averaging through a mathematical manipulation whose side-effect is to drastically reduce computational and storage requirements. Aside from these features that represent a considerable advantage for practical applications, the classical ESPRIT technique has the same average spatial resolution and precision as MUSIC.

Subspace-based techniques are generally seen as a trade-off between computational complexity, high resolution, and robustness to low SNR. While MUSIC does still stand as a standard for educational purpose, its performance is poor when compared to its peers such as Root-MUSIC [47], ESPRIT, IQML (Iterative Quadratic Maximum Likelihood) [48], or Root-WSF (Weighted Subspace Fitting) [49]. While not discussed in the remaining of this section, these techniques have been envisaged and their performance assessed for the antenna array introduced in Chapter 5.

3.5 Maximum Likelihood Techniques

Also referred to as *parametric beamformers*, Maximum Likelihood (ML) based techniques were some of the first investigated for DOA estimation [50]. ML techniques can usually be interpreted as a beamforming optimized by a parametric modelization of

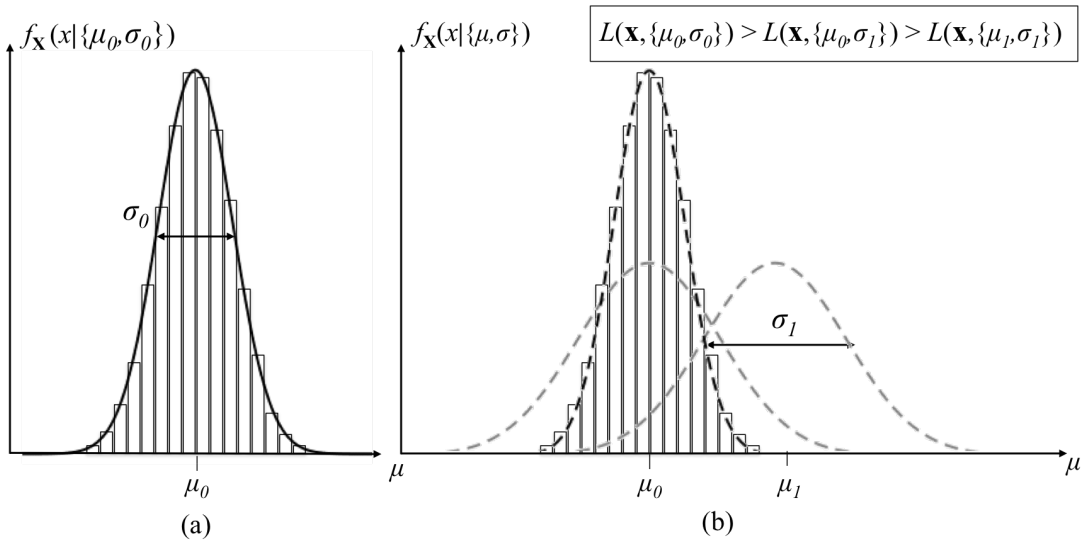


Figure 3.6: Illustration of a ML estimation. Let us assume that the pdf $f_{\mathbf{x}}(x|\theta_0)$ follows a normal distribution of true mean μ_0 and variance σ_0^2 with $\theta_0 = \{\mu_0, \sigma_0\}$. (a) Original distribution $f_{\mathbf{x}}(x|\theta_0)$ and observations $[x_1, x_2, ..., x_N]$. (b) ML estimation of μ_0 and σ_0 by modifying μ and σ in $f_{\mathbf{x}}$. As $\theta = \{\mu, \sigma\}$ is modified towards the true value $\theta_0 = \{\mu_0, \sigma_0\}$ that produced the observations $[x_1, x_2, ..., x_N]$, the value of $L(\mathbf{x}, \{\mu, \sigma\}) = \prod_{n=1}^{N} f_{\mathbf{x}}(x_n|\{\mu, \sigma\})$ increases until it reaches a maximum.

the probability density function (pdf) of received data vector $\mathbf{x}(\mathbf{t})$. The ML principle can be resumed as follow:

> If one assumes an observed dataset to be issued from a specific pdf, and that this pdf shape is itself assumed to depend on a parameter θ, fixed at an unknown θ_0 during the aforesaid observation, the value of θ that shapes the pdf to fit the observed data is likely to be θ_0.

For a more graphical illustration, let us assume an unknown DOA θ_0 to be estimated through observations $\mathbf{x} = [x_1, x_2, ..., x_N]$. Let $f_{\mathbf{x}}(x|\theta)$ design the pdf of these observations (for DOA value θ) and $L(\mathbf{x}, \theta)$ its associated likelihood function, defined as

$$L(\mathbf{x}, \theta) = \prod_{n=1}^{N} f_{\mathbf{x}}(x|\theta) \tag{3.16}$$

As illustrated in Figure 3.6, (3.16) has its maximum in $\theta = \theta_0$.

The ML approach known as the *least square method* attempts to minimize the difference between the received signal $\mathbf{x}(t)$ and a parametric estimate $\mathbf{A}(\hat{\theta})\hat{\mathbf{s}}(t)$, assuming that the original $\mathbf{x}(t) - \mathbf{A}(\theta)\mathbf{s}(t) = \mathbf{n}(t)$ in (3.3) follows a normal distribution. In most

implementations the estimated parameters are reduced those of a single source, i.e. $\{\hat{\theta}_k, \hat{\mathbf{s}}_k\}$ instead of $\{\hat{\theta}_1, ..., \hat{\theta}_K, \hat{\mathbf{s}}_1, ..., \hat{\mathbf{s}}_K\}$, gathering the signals from other sources in the interference vector $\mathbf{n}(t)$. As their is no major subtleties in the mathematics related to basic ML estimation, let us simply assume that

$$\min_{\hat{\theta},\hat{\mathbf{s}}} \left[\|\mathbf{x}(t) - \mathbf{A}(\hat{\theta})\hat{\mathbf{s}}(t)\|^2 \right]$$

is equivalent to

$$\max_{\hat{\theta}} \left[trace \left\{ \mathbf{P}_A(\hat{\theta})\mathbf{R}_x \right\} \right] \tag{3.17}$$

where $\mathbf{P}_A(\hat{\theta})$ is the matrix of the space spanned by the columns of $\mathbf{A}(\hat{\theta})$ defined as

$$\mathbf{P}_A(\hat{\theta}) = \mathbf{A}(\hat{\theta}) \left(\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta}) \right)^{-1} \mathbf{A}^H(\hat{\theta})$$

Based on (3.17), the spatial power spectrum $P_{leastsquare}(\theta)$ of the ML least square method is then defined as

$$P_{leastsquare}(\theta) = trace \left\{ \mathbf{P}_A(\hat{\theta})\mathbf{R}_x \right\}$$

To evaluate the efficiency of the ML least square estimation technique, the simulation scenario introduced in Section 3.2 is applied to the DOA estimation of 3 uncorrelated sources of respective DOAs $\{\theta_1, \theta_2\, \theta_3\} = \{-40^\circ, 10^\circ, 25^\circ\}$. The resulting spatial power spectrum is reported in Figure 3.7, suggesting a resolution equivalent to the one observed with the Bartlett beamformer with so few samples in the observation vector $\mathbf{x}(t)$.

In terms of optimum performance, ML techniques are usually superior to other estimators (especially under low SNR) while being known for their high computational complexity. Due to their relatively generic formulation, ML techniques are interesting both in terms of didactic and cross-domain comparison studies.

3.6 A Word on Antenna Array Geometry

As for the review of DOA estimation techniques, an exhaustive listing of every possible array geometry would be irrelevant here. The work reported in [51] is a good introduction on the subject's complexity.

Seen from a geometrical perspective, the main characteristics of an array geometry are its dimension (1,2, or 3D)[5] and the distribution of its elements (uniform, non-

[5]The term array dimension refers to the minimum dimension of a cartesian space required to describe its antennas disposition.

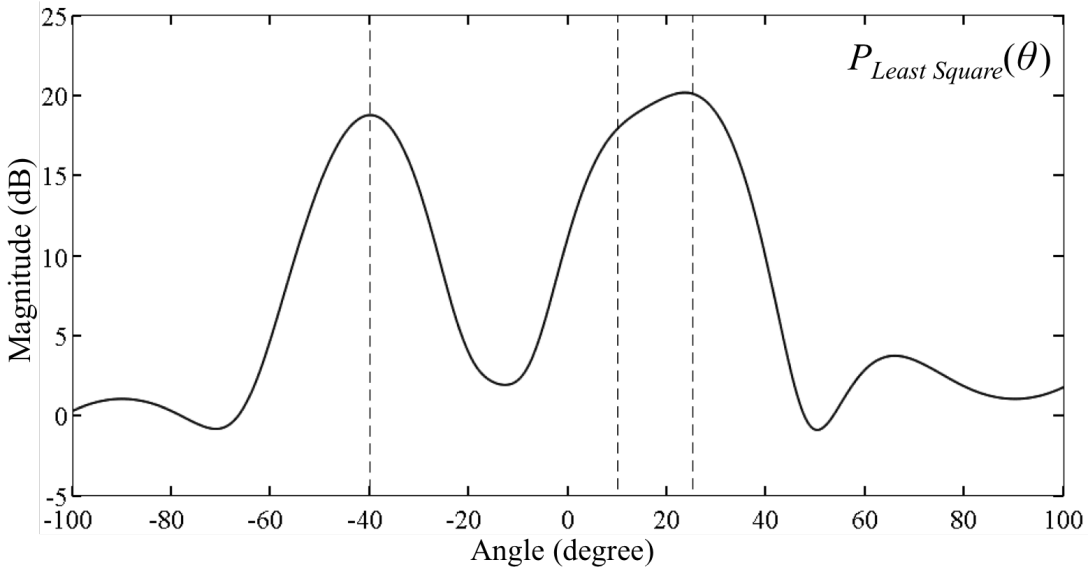


Figure 3.7: Plot of the spatial power spectrum $P_{leastsquare}(\theta)$ associated to the ML least square estimation technique for DOAs $\{\theta_1, \theta_2\, \theta_3\} = \{-40^\circ, 10^\circ, 25^\circ\}$. The estimate of the left hand source DOA at -40° is correct but sources at 10° and 25° are not *resolved*, much alike the Bartlett beamformer.

uniform). To quote but a few, the most frequently encountered geometries are

- Uniform Linear (1D), Circular (2D), Rectangular Array (2D)
- L or T-shaped array (2D)
- Uniform Spherical, Uniform Cubic Array (3D)

Seen from a DOA estimation perspective, these characteristics become the dimension of the DOA the array can estimate (1D for azimuth or 2D for azimuth and elevation), its resolution and the isotropy or non-isotropy of its resolution. Some estimation techniques can be applied only to certain geometries, like the classical ESPRIT that requires shift-invariant geometries. This is not usually a limiting factor regarding the choice of the array geometry as there exists adaptations of most of the more efficient, performant, or robust techniques to most array geometries [52–54].

The major issue when designing a DOA estimator's array usually concerns its number of antennas. Too few elements will significantly limit the optimum performance of the DOA estimator. Arrays composed of e.g. less than 5 antennas are generally not compatible with high resolution techniques or pre-processing schemes required to operate estimations under low SNR or multi-path propagation conditions [55].

Besides an increase in computational complexity, the main issue of an array composed of a large amount of antennas is its size. For most geometries, the size of an array

is defined by its number of elements and its average inter-element "step-size" (c.f. Figure 3.1). The inter-element spacing is generally proportional to the signal's wavelength, fixed at $\lambda/2$ in most implementations. Too short an inter-element spacing compared to this value will make the phase difference between adjacent antenna elements too small to be measured, even for high resolution algorithms. This "oversampling" array would not yield much better estimations than an array of equal size with less elements spaced by $\lambda/2$. On the other hand, geometries with an inter-element spacing greater than $\lambda/2$ will be affected by *spatial aliasing*, i.e. will admit two or more solution to a given DOA estimation problem. By analogy with temporal sampling, inter-element spacing is often defined around or below half a wavelength to comply with the *spatial Nyquist-Shannon sampling theorem*.

3.7 Conclusion

This chapter presented the fundamentals of DOA estimator implementation, from DOA estimation techniques to array geometry design. Beamforming, subspace-based decomposition, and ML-based estimation techniques have been introduced and their resolution illustrated through a case-study estimation scenario based on a ULA geometry. Table 3.7 reports the obtained performance for each of these techniques, putting forward the subspace-based methods as a trade-off between resolution and computational complexity. The comparative study in [39] provides a likewise table, extended to the main DOA estimation techniques investigated over the last two decades.

Each technique has its assets regarding specific DOA estimation scenarios: low complexity, high resolution, robustness to low SNR or coherent signals, etc. Designing a Direction Finder requires a good understanding of the propagation conditions it is intended to operate in, along with the requirements it has to fulfill. Needless to say, a design yelding a 1° accuracy but unable to deal with coherent signals or requiring 5 minutes to produce an estimation would not be suitable the context of the targeted search and rescue application of the current study.

DOA estimation remains an active research domain, spurred by the ever present need for more robust and reliable localization solutions for civilian and military applications. In view of the topics under study in recent publications (see e.g. [36,39,51,56–60]), current research in the field focuses on DOA estimation:

- ➙ With non-uniform arrays, three-dimensional arrays
- ➙ With sparse arrays based on polarized antennas
- ➙ Of wideband and moving signal sources
- ➙ With phase excitation and virtual antenna based estimators

Table 3.1: Characteristics summary of DOA estimation techniques. The "Resolution" value have been obtained through Monte Carlo simulations (100 iterations) based on the scenario of Section 3.2 with a resolution threshold of 85% (i.e. the technique dissociated two DOAs spaced by "Resolution" in more than 85% of the iterations). The "Complexity" is an arbitrary evaluation based on a relative comparison of each techniques on a personal laptop/implementation, see [39] for a more thorough comparison.

DOA method	Resolution**	Coherent signals	Complexity **
Bartlett	18°	no	low
Capon	13°	no	medium
MUSIC	6°	no	medium
ESPRIT	6°	yes	low
ML least square	18°	yes	high

* Best average resolution achieved for the scenario introduced in Section 3.2
** Relative evaluation based on average CPU, RAM, and ROM consumption.

- In the presence of interferences or mutual coupling among elements
- With unknown or erroneous geometry model

Each of these topics attempts to widen the scope of DOA estimation application, e.g. with ever more detailed assumptions on the received signal and modeling of the propagation channel to improve the performance and robustness of these estimations. The concept of "virtual antennas" was at first introduced to loosen the constraint imposed on array geometries by ESPRIT-like techniques (see end of Section 3.4). It has since been employed to further dissociate both array and signal processing layers [35], designing interpolation and projection-based methods that shaped the received data to be compliant with the requirements of any estimation technique [61]. Finally, a trend recently appeared advocating the assessment of DOA estimation techniques on physical prototypes rather than on simulated architectures [62]. The former eventually allows one to thoroughly evaluate the performance of a given DOA estimator under real conditions. This phenomenon is not unconnected with the recent developments in Software Defined Radio technologies, discussed in Chapter 11.

Chapter 4

Performance Evaluation Tools of a DOA Estimator

Contents

Following the overview of the main Direction Of Arrival (DOA) estimation techniques and antenna array design presented in Chapter 3, this chapter introduces some evaluation tools to assess the performance of a DOA estimator[1]. This evaluation concerns the estimator's accuracy, its resolution, and its rate of convergence when confronted with different noise levels, multi-path propagation, multi-sources conditions, etc.

4.1 Introduction

The evaluation of a DOA estimator concerns the characterization of its capacity to deliver valid estimations when faced with various propagation scenarios. As for any evaluation, it implies a set of metrics to objectively assess the DOA estimator performance both across design versions and propagation conditions. Table 4.1 lists the metrics discussed in this section and the main parameters of both DOA estimator design and propagation model. The remaining of this section introduces concepts related to the accuracy and variance of an estimator.

[1] A DOA estimator designs the combination of an antenna array and a DOA estimation technique.

Table 4.1: Evaluation metrics of a DOA estimator, parameters of the DOA estimator, and parameters of the propagation scenario. A typical evaluation implies to modify the parameters of the DOA estimator to observe the impact of these modifications on the evaluation metrics across propagation scenarios.

Performance evaluation metrics of a DOA estimator
Accuracy (Mean Squared Error)
Variance
Resolution (minimum angle between two DOAs to avoid merging)
Convergence rate (number of samples required for a given variance and accuracy)

DOA estimator design parameters	
Antenna array geometry	ULA, UCA, etc.
Antennas properties	Directional, Omni, Polarized, etc.
Number of antennas	2, 3, ..., M
DOA estimation technique	Beamforming, MUSIC, etc.
Number of snapshots per estimation	1, 2, ..., N

Propagation scenario parameters	
Number of sources	1, 2, ..., K
Angle between sources DOAs	10°, 20°, etc.
Noise level	SNR = -10dB, 12dB, etc.
Multi-path propagation, interferences	correlated / uncorrelated sources

Given an estimator $T_\theta(\mathbf{x})$ of the parameter θ based on observations $\mathbf{x}$, the characterization of T_θ accuracy is based on the evaluation of the variance and the Mean Squared Error (MSE) of its estimations. MSE is defined as the expected value (probability-weighted average, over all samples) of the squared errors

$$\mathrm{MSE}\left(T_\theta\right) = E\left\{\left(T_\theta(\mathbf{x}) - \theta\right)^2\right\}$$

and the variance as the expected value of the squared sampling deviations

$$\mathrm{Var}(T_\theta) = E\left\{\left(T_\theta(\mathbf{x}) - E\left\{T_\theta(\mathbf{x})\right\}\right)^2\right\}$$

where $E\{\cdot\}$ denotes the statistical expectation. MSE and Var respectively concern the precision and the clustering of the estimations. If the parameter θ was the center of a target, and arrows thrown at it were estimates, a high MSE would suggest a high average distance from arrows to target center and a high Var that the arrows were rather dispersed (i.e. non-clustered) on the target.

As most DOA estimation techniques are based on unbiased estimators, typical characterizations are generally limited to the study of the variance's evolution regarding the number of antennas, of snapshots, the SNR level, etc.[2] Once a DOA estimator T_θ yields acceptable results in terms of estimations' variance, the evaluation will typically focus on its resolution or on its robustness to correlated signals, modifying the propagation scenario until it reflects T_θ's foreseen operational reality.

The following subsections introduce two different approaches to evaluate the performance of a DOA estimator. The first is based on the Cramer-Rao Bound (CRB) issued from the information theory. This analytical tool allows quick assessment of the optimal accuracy of a given array geometry. The second approach is based on a Monte Carlo design: average performance observation over iterative simulations in various propagation conditions. Two different propagation models are presented, respectively used to assess the statistical performance of a DOA estimator design and its punctual behavior in specific search environments.

4.2 Theoretical Evaluation, the Cramer-Rao Bound

The Cramer-Rao Bound (CRB) [63], related to the concept of Fisher Information [64], represents the theoretical limit of the average precision with which a DOA can be estimated for a given antenna array. Once derived for a given array geometry, the CRB is a powerful evaluation tool that gives instant access to what can be expected from any unbiased estimator applied to this geometry, providing a lower bound of its variance as a function of DOA, number of antennas, SNR, etc. The CRB allows the assessment of the optimum performance of an antenna array regardless of any DOA estimation technique. This section introduces the Fisher Information used to define the CRB and presents its formulation for the M-element Uniform Linear Array (ULA) introduced in Section 3.2, Chapter 3.

Let us assume the estimation of DOA θ_0 through observations $\mathbf{x} = [x_1, x_2, ..., x_N]$ of probability density function (pdf) $f_\mathbf{x}(x|\theta)$ and likelihood function $L(\mathbf{x}, \theta)$, as defined in Equation (3.16) Section 3.5. The concept behind any CRB-based optimization can be reduced to the observation:

The Fisher Information on θ_0, contained in observations $[x_1, x_2, ..., x_N]$, is proportional to the curvature of $L(\mathbf{x}, \theta)$ for $\theta = \theta_0$.

To simplify mathematical manipulations, the computation of the Fisher Information

[2] The statistical expectation of an unbiased estimator is equal to the true value of the estimated parameter, i.e. $E\{T_\theta\} = \theta$ which involves $\mathrm{Var}(T_\theta) = \mathrm{MSE}(T_\theta)$.

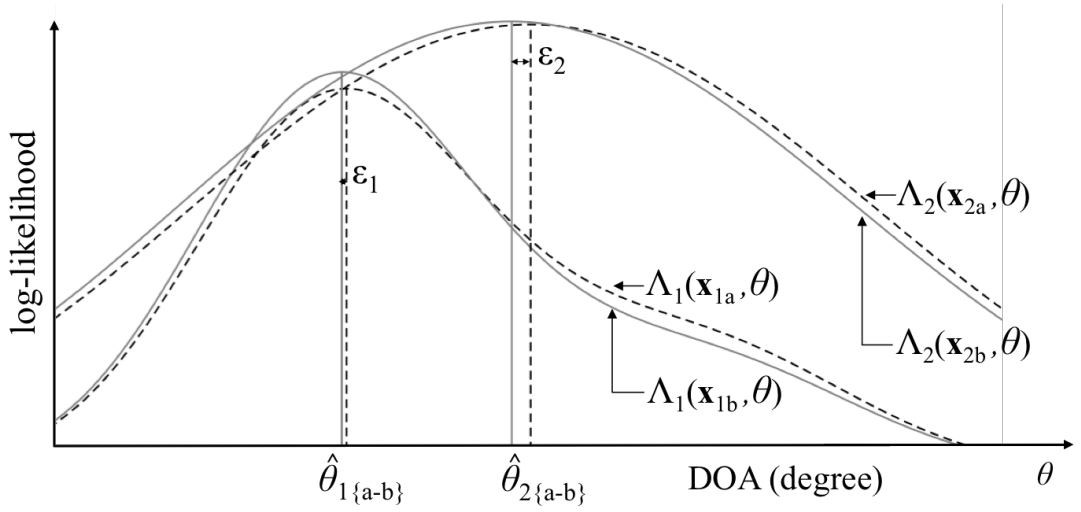


Figure 4.1: Illustration of the Fisher Information principle related to the curvature of the log-likelihood. Let us assume two sets of observations $\mathbf{x}_{1a}$ and $\mathbf{x}_{2a}$ issued from pdfs $f_{\mathbf{x}1}(x|\theta_1)$ and $f_{\mathbf{x}2}(x|\theta_2)$. Log-likelihood functions $\Lambda_1(\mathbf{x}_{1a},\theta)$ and $\Lambda_2(\mathbf{x}_{2a},\theta)$ are computed to estimate θ_1 and θ_2, yielding respective maximums in $\hat{\theta}_{1a}$ and $\hat{\theta}_{2a}$. For a slightly different set of observations $\mathbf{x}_{1b}$ and $\mathbf{x}_{2b}$ (issued from the same pdfs), the log-likelihood functions will e.g. have their maximum in $\hat{\theta}_{1b}$ and $\hat{\theta}_{2b}$. Due to the difference in curvature between Λ_1 and Λ_2 around respective DOAs θ_1 and θ_2, the shift ε_1 between $\hat{\theta}_{1a}$ and $\hat{\theta}_{1b}$ is lower than its counterpart $\varepsilon_2 = \hat{\theta}_{2a} - \hat{\theta}_{2b}$. A detailed derivation based on Λ_1, Λ_2, and (4.1) would likely yield $I(\theta_1) > I_2(\theta_2)$.

is generally based on the *log-likelihood* function $\Lambda(\mathbf{x},\theta)$ rather than the likelyhood $L(\mathbf{x},\theta)$, defined as

$$\Lambda(\mathbf{x},\theta) = \log\big(L(\mathbf{x},\theta)\big) = \sum_{n=1}^{N} \log\big(f_{\mathbf{x}}(x_n|\theta)\big)$$

where the introduction of the logarithm allows the manipulation of a sum of pdfs instead of a product of pdfs. As the following results only depend on the relative variations of the likelihood function, both likelihood or log-likelihood based analysis are equivalent since log is an increasing function[3]. Figure 4.1 illustrates the concept of Fisher Information and its relation to the curvature of the log-likelihood function in θ, i.e. its second derivative with respect to θ according to differential geometry. As $\Lambda(\mathbf{x},\theta)$ inherited the random nature of $\mathbf{x}$, the definition of the Fisher Information, denoted $I_N(\theta)$, involves the statistical expectation of the log-likelihood[4]:

[3] or monotone function, a function between ordered sets that preserves the given order.

[4] (4.1) is based on the opposite of the second derivative of $\Lambda(\mathbf{x},\theta)$ since $\Lambda(\mathbf{x},\theta_0)$ is a local maximum.

$$I_N(\theta) = -E\left\{\frac{\delta^2}{\delta\theta^2}\Lambda(\mathbf{x},\theta)\right\} \tag{4.1}$$

where $I_N(\theta_0)$ increases as the lobe of $\Lambda(\mathbf{x},\theta)$ in θ_0 gets "sharper". The Cramer-Rao theorem then states that the variance expected from any unbiased estimator $T_\theta(\mathbf{x})$ of the parameter θ based on observations $\mathbf{x}$ cannot get lower than the inverse of the Fisher Information contained in $\mathbf{x}$, written as

$$\mathrm{Var}\left(T_\theta(\mathbf{x})\right) \geqslant \frac{1}{I_N(\theta)} = \frac{-1}{E\left\{\frac{\delta^2}{\delta\theta^2}\Lambda(\mathbf{x},\theta)\right\}} = \mathrm{CRB}_\theta(\theta) \tag{4.2}$$

The notation CRB_θ indicates that the bound concerns the variance over the estimation of θ. The CRB formulation for a 2D DOA estimation problem will involve the Fisher Information Matrix (FIM), the matrix format of the Fisher Information based on the Hessian matrix of the log-likelihood. Defining the FIM as $\mathbf{F}(\mathbf{\Theta}) = -E[\frac{\partial^2\Lambda(\mathbf{\Theta})}{\partial\mathbf{\Theta}^2}]$ for the 2D DOA $\mathbf{\Theta} = \{\theta,\varphi\}$, the Cramer-Rao theorem extended to its matrix format states that

$$\mathbf{R}(T_{\mathbf{\Theta}}) - \mathbf{F}(\mathbf{\Theta})^{-1} \tag{4.3}$$

is a positive definite matrix, where $\mathbf{R}(T_{\mathbf{\Theta}})$ is the covariance matrix of $T_{\mathbf{\Theta}}$, unbiased estimator of $\mathbf{\Theta}$. This formulation of the FIM is further used in Chapter 5 to derive both CRB_θ and CRB_φ for a spherical array configuration.

Based on the assumptions and signal model presented in Section 3.2, the CRB related to the estimation of a single DOA θ with a M-element ULA can be expressed as

$$\mathrm{CRB}_\theta(\theta_0) = \frac{6}{\left(\frac{2\pi d}{\lambda}\right)^2 M\left(M^2-1\right)\mathrm{SNR}\cos^2\theta_0} \tag{4.4}$$

where $\mathrm{SNR} = \sigma_s^2/\sigma_n$ with σ_s^2 and σ_n^2 the respective signal and noise variances. The derivation, detailed in [35], is based on the assumption that the difference $\mathbf{x}(t) - \mathbf{A}(\theta)s(t) = \mathbf{n}(t)$ follows a normal distribution. Figure 4.2 presents a typical use of the CRB to assess the optimal performance of a given array geometry, illustrating the variation of CRB_θ in (4.4) with respect to number of antennas, SNR, and source DOA. As expected, the CRB decreases as the SNR and number of antenna elements increase. According to Figure 4.2c, ULA arrays are more accurate at estimating DOAs perpendicular to the line formed by their antennas. This result is generalized for 3D geometries in [65] which are more accurate at estimating DOAs perpendicular to what is referred to as their "principal axes of inertia".

Thus defined, $I_N(\theta)$ is positive and increases with the curvature of $\Lambda(\mathbf{x},\theta)$ in θ_0.

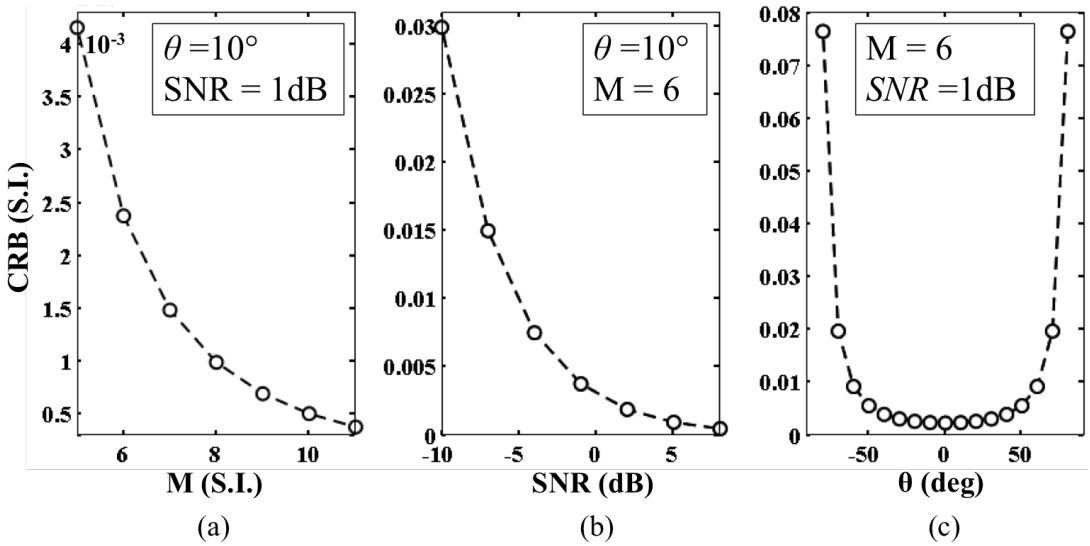


Figure 4.2: Variation of CRB_θ in Equation (4.4) for a ULA of inter-element spacing $d = \lambda/2$ as a function of (a) the number of antennas M, (b) the SNR, and (c) the source DOA θ.

4.3 Simulation-Based Evaluation

This section illustrates how to evaluate the performances of a DOA estimator through simulation. First the Monte Carlo method is introduced and applied to a M-element ULA, yielding results similar to the application of the CRB in the previous section. Two different propagation models are then discussed, respectively based on statistical and raytracing propagation simulation. A short case-study is introduced to understand how these models allow one to extend the evaluation of a DOA estimator and assess its behavior when faced with realistic propagation conditions.

4.3.1 Statistical Evaluation, the Monte Carlo Method

The Monte Carlo method requires the running of the same simulation many times over in order to obtain the distribution of an unknown probabilistic entity [66]. Applied to a DOA estimator T_θ, it allows the estimation of its statistical expectation and variance under specific propagation conditions. Figure 4.3 illustrates the typical result of a Monte Carlo simulation on a ULA array for the exact same propagation scenario as in the CRB illustration in Figure 4.2. The estimation is based on the Bartlett beamformer. As this estimation technique is not asymptotically efficient[5], the variance

[5]For a given array geometry, a DOA estimation technique that produces estimations with a variance as low as the CRB associated to this array for an infinite number of observations is said to be

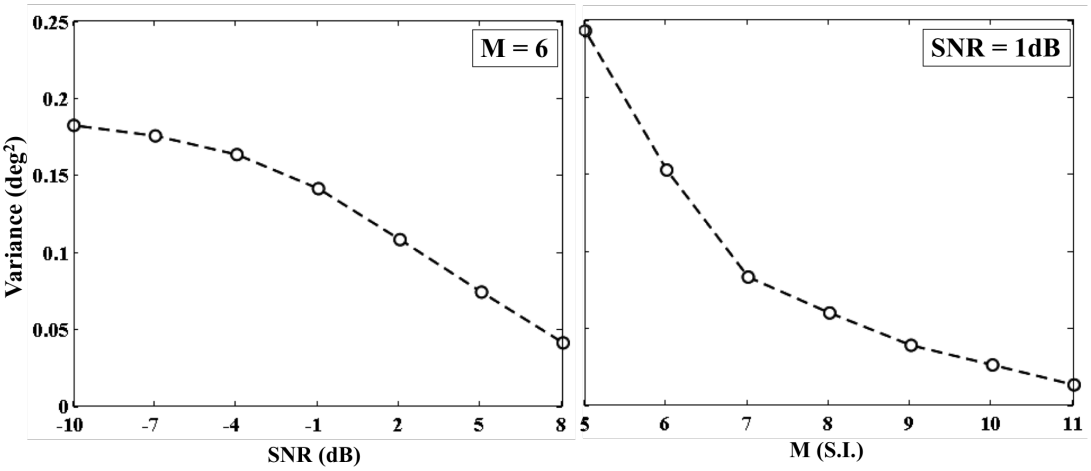


Figure 4.3: Variance of DOA estimations for a M-element ULA of inter-element step-size $d = \lambda/2$ coupled with the Bartlett beamformer introduced in Section 3.3 as a function of (a) the SNR and (b) the number of antennas M. Aside from the SNR values, simulations are based on the propagation scenario of Section 3.2 for a single source of DOA $\theta = 10^{\circ}$. Each variance value ($\circ$) is computed for 400 DOA estimations.

of its estimations is larger than the CRB. Chapter 5 presents a similar comparison between the CRB and the variance of estimations based on an asymptotically efficient estimation technique.

Compared to the CRB, this method does not involve any derivation, is valid for any propagation scenario, and allows the assessment of the performance of a specific DOA estimation technique. Its main issue is that it requires processing a large number of DOA estimations for a fixed set of parameters to produce statistically relevant values.

4.3.2 Robustness Evaluation in Realistic Propagation Scenarios

Up to this point, the DOA estimation techniques applied on the simple propagation scenario introduced in Section 3.2 have produced acceptable results. As one might expect, these results will not be as good for a DOA estimator operating in more complex propagation conditions, i.e. when faced with multi-path. While it is quite possible to manually simulate multi-path propagation, the purpose of the propagation models described in this section is that they allow one to do so coherently with a specific propagation environment, e.g. that reflects the one foreseen for the DOA estimator application.

asymptotically efficient, see [67] for a detailed discussion.

Until recently, the main drawback related to the use of realistic propagation models in small and medium sized projects was related to their cost. To support quality and competitiveness in the development of next generation radio interfaces, the European Commission funded the WINNER international research project [68]. WINNER's objective was to gather scientists and companies to design an open source propagation model that would enable free and fast assessment of any 3G-related radio interface (from 2 to 6 GHz with up to 100 MHz bandwidth). Based on both literature and extensive measurement campaigns for various propagation scenarios, they created WINNER II[6], a Matlab-based open source statistical propagation model [70].

Amongst other things, WINNER II supports multi-antenna, multi-path, polarization and 3D propagation scenarios. Rather than simulating the complete emitter-to-receiver propagation, computation involve a statistical generation of the channel parameters, based on probability density functions specific to the considered propagation scenario. The associated light computation complexity makes the WINNER II model perfectly fit for fast Monte Carlo simulations. While the WINNER II model has been designed for Radio Frequency (RF) propagation from 2 to 6 GHz, most of its hypothesis related to DOA estimation (DOAs, scattering, etc.) hold for propagation in the GSM-900 and DCS-1800 bands (0.9 and 1.8 GHz see Table 2.2). The only limitation of this model regarding the simulation of GSM propagation would be related its path loss coefficients, detailed in [70], no longer valid for carrier frequencies below 2 GHz. This has little impact on the evaluations presented in this manuscript as WINNER II is used for relative comparisons between DOA estimation techniques where the same path loss is applied in every simulation.

Table 4.2 reports some of the propagation scenarios available in WINNER II. Figure 4.4 illustrates a punctual DOA estimation based on a 4-element UCA coupled with a MUSIC estimator, showing the impact of typical indoor/outdoor multi-path propagation on DOA estimation.

Table 4.2: Some of the available propagation scenarios in the WINNER II model as presented in [70]. Scenarios like A1, B4 or C1 are particularly relevant for the considered search and rescue application.

Scenario	A1	A2	B1	B3	B4	C1
Definition	Indoor office	Indoor to outdoor	Urban microcell	Large indoor hall	Outdoor to indoor	Suburban

[6]WINNER II follows the 3GPP SCM and IMT-Advanced geometric channel models, see [69] for a detailed comparison.

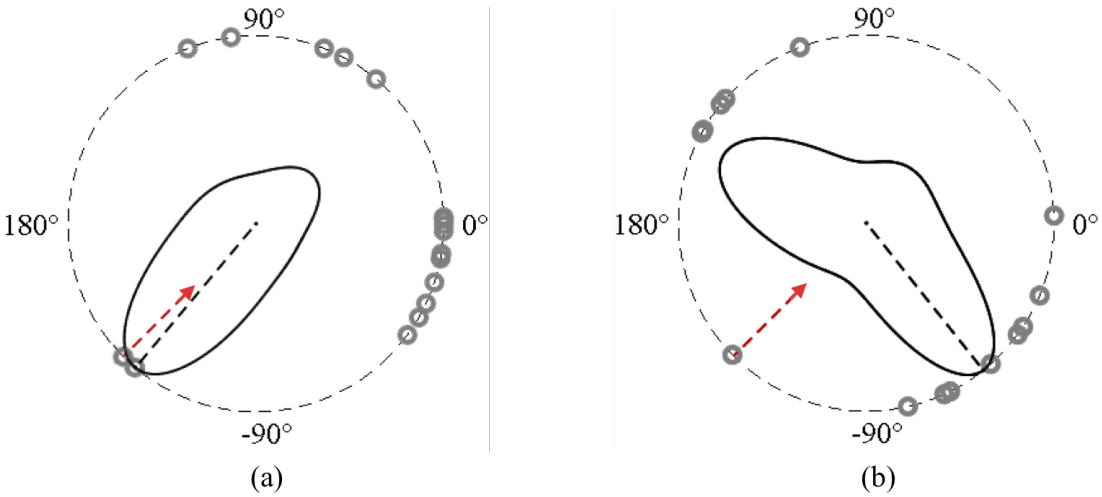


Figure 4.4: MUSIC spectrum of a 4-element UCA for two different WINNER II scenarios (a) B1 (urban microcell) and (b) B3 (large indoor hall). Both scenarios assume that the single emitter is in line of sight, positioned respectively at 70 m and 20 m from the receiver. Red arrow is the true DOA, grey circles represent the signals impinging on the array due to multi-path propagation. Black outline and dotted line are the MUSIC spatial power spectrum and estimated DOA respectively. Both simulations assume noiseless conditions.

4.3.3 Behavioral Evaluation, DOA Estimation for Navigation

As the considered application involves the reliance on DOA estimations to navigate from point A to point B, the evaluation of the DOA estimator cannot rely on statistical tests only. It must somehow take into account its dynamic "behavior" during the navigation. The model presented in this section belongs to a class referred to as *geometric* propagation models, based on a raytracing simulation in a tangible environment. Geometric models allow one to observe the impact of the elements of the environment (walls, windows, etc.) onto both the RF propagation and the DOA estimation. The objective of the simulations conducted with the raytracing model described below is to facilitates the understanding of how the estimations will behave during the navigation as the DOA estimator is confronted with a constantly evolving propagation scenario.

As raytracing models usually raise the same financial issue as statistical models, the IlmProp raytracing model [71] released in 2006 by the Ilmenau University is a most welcome element to complete the assessment toolbox proposed in this chapter. This open source model allows one to define the geometry of a propagation scenario, manually positioning emitter, receiver, obstacles, scatterers, etc. Limited to 2D scenarios, the current implementation does not support realistic obstacle's absorption, by default opaque regarding RF signals. This will eventually restrict the complexity of the propagation scenarios that can be assessed with this model.

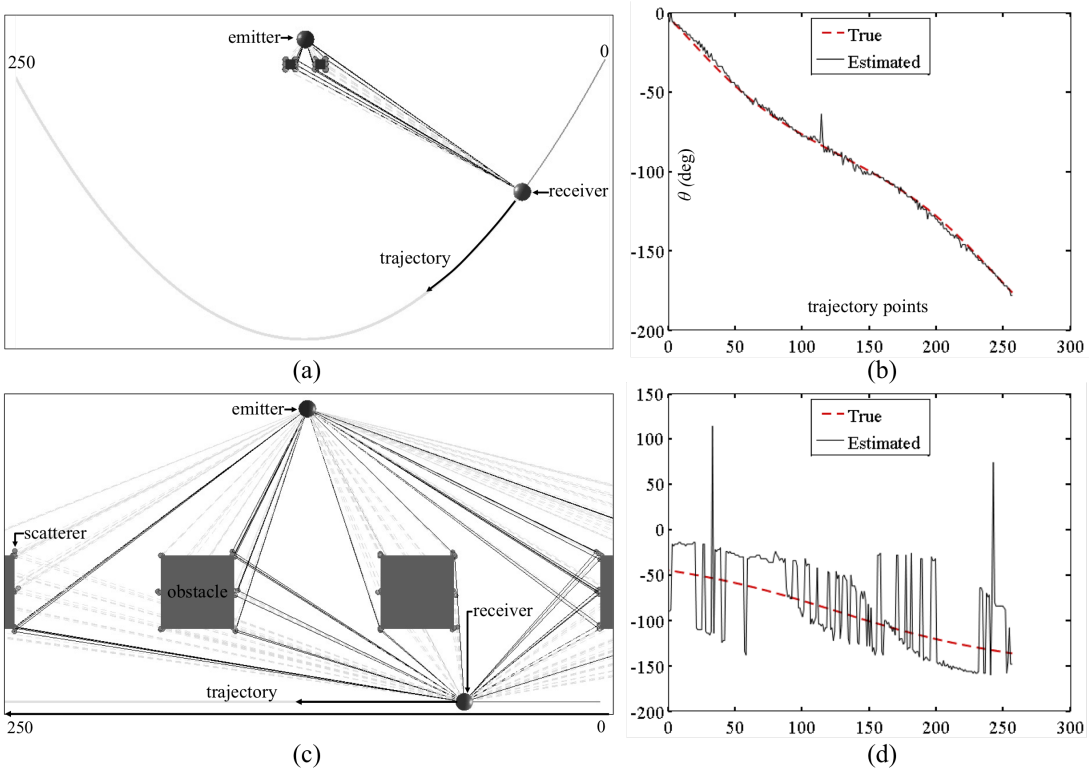


Figure 4.5: Illustration of a typical simulation using the IlmProp raytracing model. Emitter, obstacle, and scatterer positions are defined prior to the simulation, along with the antenna array geometry and the receiver (DOA estimator) trajectory. IlmProp outputs an array of impulse responses for discrete points on the trajectory, sent to the DOA estimator based on MUSIC coupled to a 8-element UCA. The estimations related to propagation scenarios (a) and (c) are shown in plots (b) and (d) respectively, across the 250 discrete positions of the receiver in each scenario. Solid black lines correspond to rays that reach the receiver. Dotted grey lines correspond to rays discarded after impinging through one of the obstacles. While transmission loss through obstacles is not supported, the free parametrization and positioning of the scatterers allows one to simulate reflections and edge diffraction. The instabilities observed in the estimations of the second scenario are related to the emitter being alternatively shadowed by the different obstacles.

Two different navigation scenarios are shown in Figure 4.5, based on a DOA estimator coupling MUSIC to a 8-element UCA. The first scenario assumes an open-field like situation, the second a more crowded environment where obstacles successively block the line-of-sight propagation path between emitter and receiver during the navigation. This simple case-study highlights the input of such a model on the general understanding of the DOA estimator behavior during the navigation, as the instabilities in its

estimations Figure 4.5d are but a consequence of the shift in propagation scenarios encountered during the navigation.

4.4 Conclusion

This Chapter presented some evaluation tools used to assess the performance of a DOA estimator. It introduced the Cramer Rao Bound, an analytical formula that allows one to quickly assess the optimal performance of any antenna array geometry in terms of estimation accuracy, prior to the selection of any specific DOA estimation technique. Two approaches to simulate the performance of a complete DOA estimator design in realistic conditions have been discussed based on two different propagation models. The first approach was based on a the WINNER II statistical model, ideal for quickly assessing the average performance of a DOA estimator in realistic propagation scenarios. The second employed the Ilmprop raytracing model to provide some insight into how a DOA estimator dynamically behaves during a navigation task.

Chapter 5

Evaluation of an Antenna Array for DOA Estimation, Case Study of a Lightweight Helmet-Mounted Design

Contents

This chapter details the implementation, evaluation, and optimization of a Direction Of Arrival (DOA) estimator designed for search and rescue applications, based on the knowledge and tools introduced in Chapters 3 and 4. An antenna array geometry is proposed that meets the size and weight constraints of a portable application. The Fisher Information Matrix (FIM) specific to this geometry is derived and its associated Cramer Rao Bounds (CRBs) are used in order to understand the impact of the array geometry on DOA estimation accuracy. The performance of the antenna array coupled with a MUSIC estimator are then assessed in realistic propagation conditions, based on the WINNER II and IlmProp propagation models introduced in Section 4.3. Part of this chapter has been published in the proceedings of the 4$^{\text{th}}$ International Congress on Ultra Modern Telecommunications and Control Systems (2012) [1].

Notations

$\mathrm{E}\{\cdot\}$	Statistical expectation
$\Re\{\cdot\}$	Real part of a complex element
$(\cdot)^H$	Conjugate transpose
$\lvert\cdot\rvert$	Determinant of a matrix or module of a complex
$\mathrm{tr}(\cdot)$	Trace operator
$\lvert\lvert\cdot\rvert\rvert$	Norm of a vector
$\langle u_1, u_2\rangle$	Scalar product between vector u_1 and u_2
I_M	$M \times M$ identity matrix
$col[u_m]$	Column vector of elements u_m, $m \in [\![1, M]\!]$
$diag(u_m)$	Diagonal matrix of elements u_m, $m \in [\![1, M]\!]$
$\delta_\xi(\cdot)$	1^{st} Partial derivative with respect to ξ
$\delta^2_{\xi,\zeta}(\cdot)$	2^{nd} Order partial derivative with respect to ξ then ζ

5.1 Introduction

As introduced in Chapter 1, the objective of the considered Direction Finder (DF) is to assist individual rescuers through search and rescue operations. Its general design has to be lightweight and compact, such as not to hinder rescuers in their movements. A sensible implementation is to backpack most of it (batteries, central unit, etc.) and position the antenna array to optimize the reception quality. A helmet-mounted array fits this requirement with the added value of producing DOA estimations related to users head orientation, a feature used later in Part II. The geometry considered in this chapter assumes an array with its antennas distributed on an hemisphere representing the helmet, illustrated in Figure 5.1. As discussed in Chapter 4, the characterization of the geometry for DOA estimation will involve the derivation of its FIM.

While extensive studies on circular and spherical antenna array optimization have already been published [72, 73], none presented a generic form of the FIM for non-uniform spherical arrays. Several papers have adressed the 3D antenna array optimization problem based on the FIM in the single source case [74,75], except for general array configurations not adapted to the architecture considered here. The approach proposed is to first focus on a simple antenna array composed of three antennas to investigate how the CRBs issued from the FIM can be decomposed into intuitive metrics. The objective of these metrics is to reduce the process of antenna positioning optimization for DOA estimation to basic geometric considerations on the general configuration of the array. Based on these considerations, the array geometry can be optimized for the estimation of certain DOAs, "focusing" its resolution in a certain direction. The considered antenna array configuration is constrained to 2 Degrees of Freedom (DoF), the DOA estimation problem limited to a single emitter to simplify the analysis, much

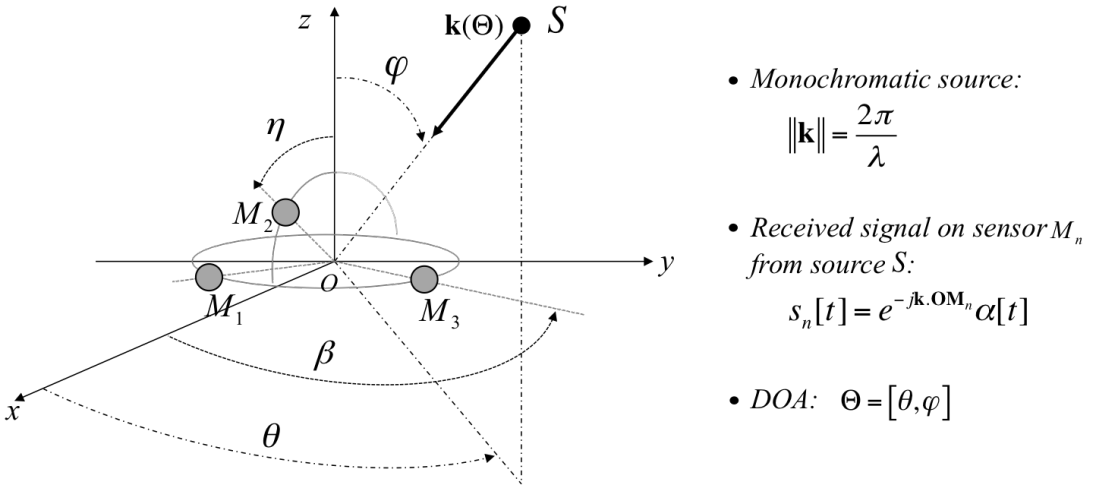


Figure 5.1: Illustration of the spherical antenna array geometry for the DOA estimation of a single source S. β and η represent the 2 DoF of the geometry configuration.

like the optimization presented in [76]. As a reminder, Table 5.1 resumes the notations used within this chapter.

5.2 Signal Model and Array Geometry

Let us assume the 3-element spherical array of Figure 5.1. Each supposedly isotropic and identical sensor M_i evolves on a sphere of radius r centered at O. $\mathbf{OM_1}$ and $\mathbf{OM_3}$ are constrained in the Oxy plane and form an angle $\pm\beta$ with Ox, $\mathbf{OM_2}$ is constrained in Oxz and forms an angle η with Oz. The array configuration is referred to as $\mathbf{\Omega} = \{\beta, \eta\}$, constrained to 2 DoF. Let S represent the single source that propagates in free space at radio wavelength λ to impinge on the array from DOA $\mathbf{\Theta} = \{\theta, \varphi\}$. For notational convenience, $\mathbf{\Theta}$ and $\mathbf{\Omega}$ will be omitted in the following derivations whenever there is no possibility of confusion.

The steering vector $\mathbf{a_\Omega(\Theta)}$ of the received signal on the array sensors can be expressed as

$$\mathbf{a_\Omega(\Theta)} = \begin{pmatrix} e^{j\frac{2\pi r}{\lambda}(sin(\varphi)cos(\theta+\beta))} \\ e^{j\frac{2\pi r}{\lambda}(sin(\varphi)cos(\theta)sin(\eta)+cos(\varphi)cos(\eta))} \\ e^{j\frac{2\pi r}{\lambda}(sin(\varphi)cos(\theta-\beta))} \end{pmatrix} \tag{5.1}$$

consequently, the vector measurement $\mathbf{x}[t]$ used to perform the DOA estimation takes the form

$$\mathbf{x}[t] = \mathbf{a}_{\mathbf{\Omega}}(\mathbf{\Theta})s[t] + \mathbf{n}[t] \tag{5.2}$$

where $s[t]$ is the signal sample and $\mathbf{n}[t]$ the vector of noise samples, both received by the array antennas at time t. Respective signal and noise variances are denoted σ_s^2 and σ_n^2.

5.3 Theoretical Evaluation and Optimization Based on the CRB

5.3.1 Formulation of the FIM and CRBs for a Spherical Geometry

As a reminder, the CRB is a *lower bound* of the covariance matrix of any unbiased estimator, first introduced by R.A. Fisher in [64], defined as the inverse of the FIM. After calculations detailed in Appendix Section 5.6, the FIM introduced in Section 4.2 for the spherical array of Figure 5.1 takes the form

$$\mathbf{F}(\mathbf{\Theta}) = \gamma_M \begin{pmatrix} \langle \mathbf{v}_\theta, \mathbf{v}_\theta \rangle_{\perp \mathbf{a}} & \langle \mathbf{v}_\theta, \mathbf{v}_\varphi \rangle_{\perp \mathbf{a}} \\ \langle \mathbf{v}_\varphi, \mathbf{v}_\theta \rangle_{\perp \mathbf{a}} & \langle \mathbf{v}_\varphi, \mathbf{v}_\varphi \rangle_{\perp \mathbf{a}} \end{pmatrix} \tag{5.3}$$

with

$$\gamma_M = 2\alpha N \frac{1}{\left(1 + \frac{1}{\alpha M}\right)} \left(\frac{2\pi r}{\lambda}\right)^2 \tag{5.4}$$

and

$$\mathbf{v}_\xi = \delta_\xi(\mathbf{U})\mathbf{a}_{\mathbf{\Omega}}(\mathbf{\Theta}) \quad \text{for } \xi \in \{\theta, \varphi\}, \quad \text{with } \mathbf{U} = \operatorname{diag}\big(u_m(\mathbf{\Omega}, \mathbf{\Theta})\big)$$

where N is the number of measured samples, $\alpha = \sigma_s^2/\sigma_n^2$ the linear SNR and M the number of array elements. $\langle ., . \rangle_{\perp \mathbf{a}}$ denotes the scalar product weighted by the orthonormal projector on the MUSIC defined noise subspace $\mathbf{\Pi}_{\perp \mathbf{a}} = \mathbf{I_M} - \tilde{\mathbf{a}}\tilde{\mathbf{a}}^H$, i.e.

$$\langle \mathbf{v}_\theta, \mathbf{v}_\varphi \rangle_{\perp \mathbf{a}} = \mathbf{v}_\theta^H \mathbf{\Pi}_{\perp \mathbf{a}} \mathbf{v}_\varphi \quad \text{with} \quad \tilde{\mathbf{a}} = \frac{\mathbf{a}}{\sqrt{\mathbf{a}^H \mathbf{a}}}$$

Except the explicit form of γ_M in (5.4), $\mathbf{F}(\mathbf{\Theta})$ in (5.3) is valid for any spherical array configuration for a single source scenario. The inversion of (5.3) being straightforward, let us consider only the upper left and lower right terms of $\mathbf{F}(\mathbf{\Theta})^{-1}$, lower bounds for the variance of respectively any θ and φ estimation, referred to as $\mathrm{CRB}_{\xi,\xi}$ for $\xi \in \{\theta, \varphi\}$. Defining the two correlation coefficients

$$\rho^2 = \frac{|\langle \mathbf{v}_\varphi, \mathbf{v}_\theta \rangle_{\perp \mathbf{a}}|^2}{\langle \mathbf{v}_\theta, \mathbf{v}_\theta \rangle_{\perp \mathbf{a}} \langle \mathbf{v}_\varphi, \mathbf{v}_\varphi \rangle_{\perp \mathbf{a}}} \quad \text{and} \quad r_\xi^2 = \frac{|\tilde{\mathbf{a}}^H \mathbf{v}_\xi|^2}{||\mathbf{v}_\xi||^2} \tag{5.5}$$

the CRB for parameter ξ adopts the simplified form

$$\mathrm{CRB}_{\xi,\xi} = \frac{1}{\gamma_M} \frac{1}{(1-\rho^2)} \frac{1}{||\mathbf{v}_\xi||^2(1-r_\xi^2)} \tag{5.6}$$

This new formulation splits each scalar $\mathrm{CRB}_{\xi,\xi}$ into three distinct components, easing further analysis of the array DOA estimation performances regarding its configuration. The two metrics ρ^2 and r_ξ^2 moreover explicitly display the numerical stability of the CRBs as both are always positive according to the Cauchy-Schwarz inequality applied on (5.5).

5.3.2 MUSIC-based CRB validation

This section illustrates how to validate the derived CRBs through MUSIC-based Monte Carlo simulations. As the MUSIC algorithm is asymptotically efficient in the single source case [67], the variance of its DOA estimations should fit the CRB for various SNR levels, DOA $\mathbf{\Theta}$, configurations $(\mathbf{\Omega}, r)$, etc.

The simulations presented concern SNR and $\mathbf{\Omega}$ variations for random values of $\mathbf{\Theta}$. DOA values that would raise trivial ambiguities due to array geometry [77] were deliberately avoided though[1]. For future applications, these ambiguities can easily be resolved by e.g. adding a fourth antenna to the geometry. The sphere radius r is assumed fixed at a quarter wavelength[2] to avoid the spatial aliasing discussed in Section 3.6. Mean Square Error (MSE) values concern 400 DOA estimations each, based on $N = 100$ snapshots.

Figure 5.2 shows the evolution of the MSE of MUSIC-based DOA estimations on θ and φ with respect to SNR, compared with the variations of both CRBs for DOA $\mathbf{\Theta} = (0, 60)$ and array configuration $\mathbf{\Omega} = (80^\circ, 10^\circ)$. Figure 5.3 shows the evolution of the MSE and both CRBs with respect to β and η for DOA $\mathbf{\Theta} = (0^\circ, 60^\circ)$ and a 15 dB SNR. The SNR is fixed to 15 dB as the MUSIC estimator appears to attain its asymptotic behavior above a threshold SNR of approximately 10 dB (see Figure 5.2).

In addition to validating the CRBs formulation in (5.6), Figure 5.3 already gives some information on the impact of array configuration onto its optimal performance. Figure 5.3(a) emphasizes the dependency of θ estimations accuracy on the Oxy array aperture (related to β). Figure 5.3(b) indicates a drop in performances in φ estimation as M_2 approaches the horizontal plane (i.e. for η close to 90°). The next section extends these observations through geometric considerations on the impact of sensors location on both ρ^2 and r_ξ^2 metrics.

[1] e.g. a source impinging from $\mathbf{\Theta} = \{0^\circ, 45^\circ\}$ on an array with configuration $\mathbf{\Omega} = \{90^\circ, 90^\circ\}$ would result in the exact same measurements than a source at $\mathbf{\Theta} = \{0^\circ, -45^\circ\}$, due to the symmetry with the plan formed by $M_1M_2M_3$.

[2] $r = \lambda/4 \approx 8$ cm for f_{GSM} around 900 MHz.

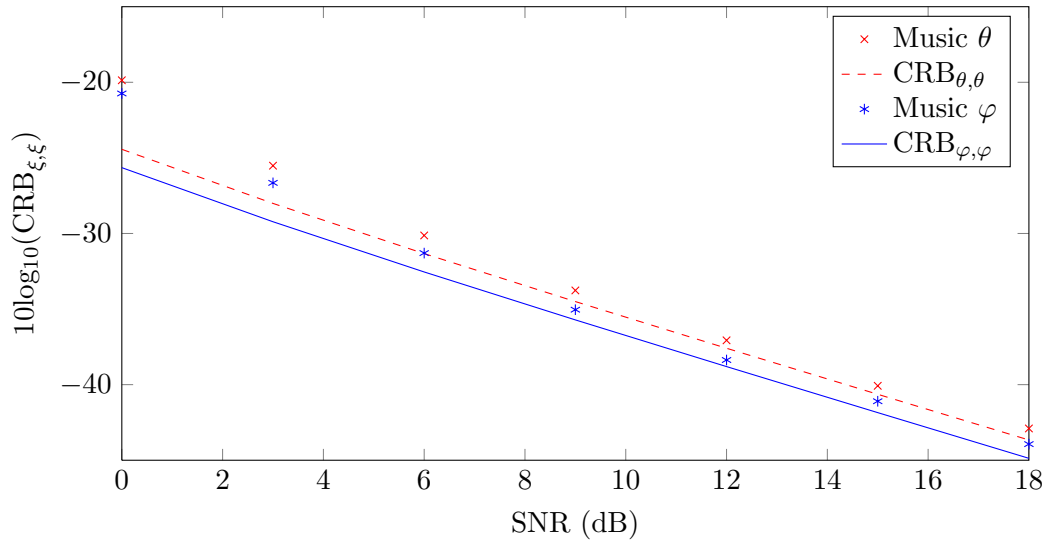
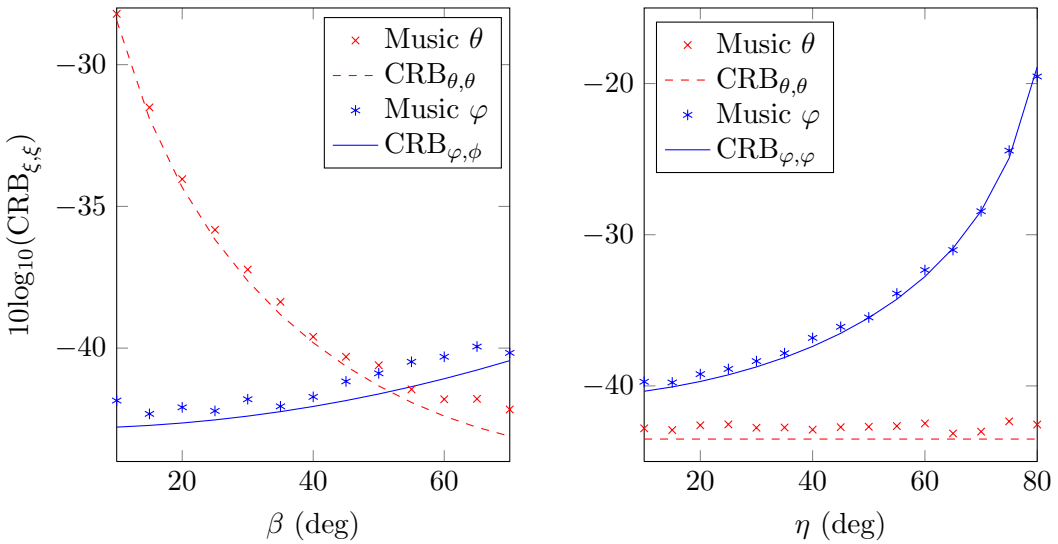


Figure 5.2: Variations of $\mathrm{CRB}_{\theta,\theta}$, $\mathrm{CRB}_{\varphi,\varphi}$, and MSEs of MUSIC estimations on DOA $\mathbf{\Theta} = (0^\circ, 60^\circ)$ with respect to SNR for a fixed array configuration $\mathbf{\Omega} = (80^\circ, 10^\circ)$.



(a) (b)

Figure 5.3: Variations of $\mathrm{CRB}_{\theta,\theta}$, $\mathrm{CRB}_{\varphi,\varphi}$, and MSEs of MUSIC estimations on DOA $\mathbf{\Theta} = (0^\circ, 60^\circ)$ with respect to (a) β (fixed $\eta = 0^\circ$) and (b) η ($\beta = 60^\circ$). SNR is fixed at 15 dB.

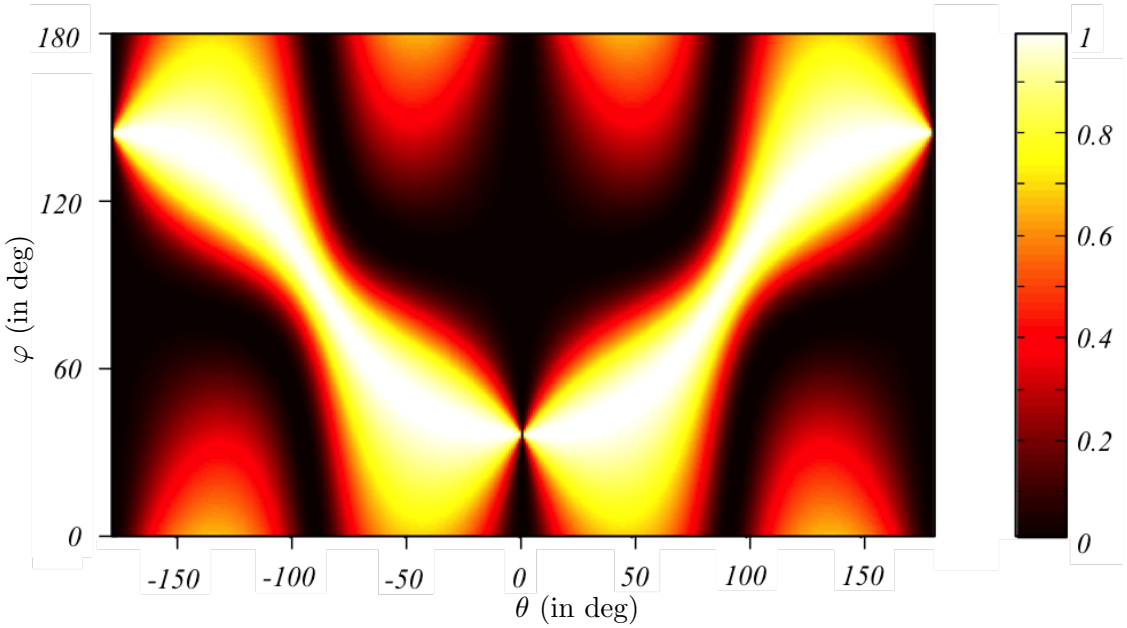


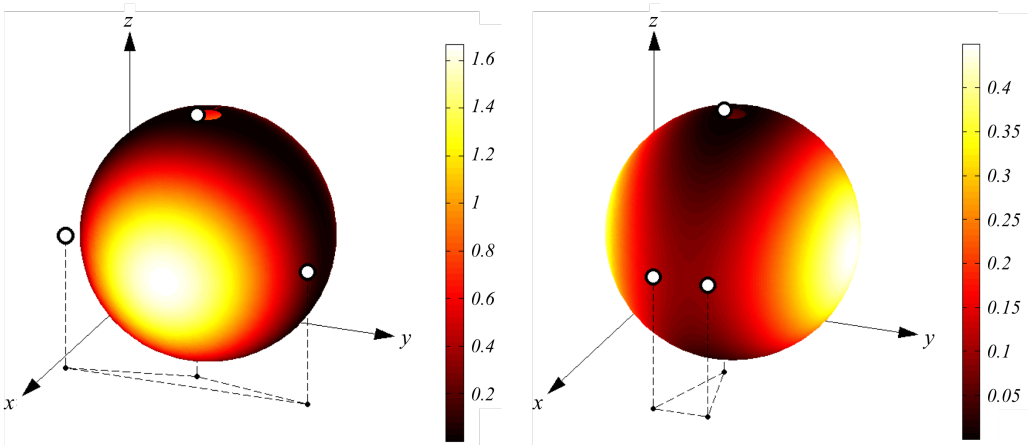
Figure 5.4: Variation of ρ^2 with respect to DOA (θ, φ) for array configuration $\mathbf{\Omega} = (80^\circ, 45^\circ)$. The least desirable DOAs are for ρ^2=1, i.e for a source in the plane $M_1M_2M_3$.

5.3.3 Array Geometry Optimization Based on the CRB

The optimization criterion presented in this section consists in minimizing ρ^2 and r_ξ^2 to reduce $\text{CRB}_{\xi,\xi}$ in (5.6). The aim is to highlight the role of these metrics in the variations of $\text{CRB}_{\xi,\xi}$ with respect to array configuration.

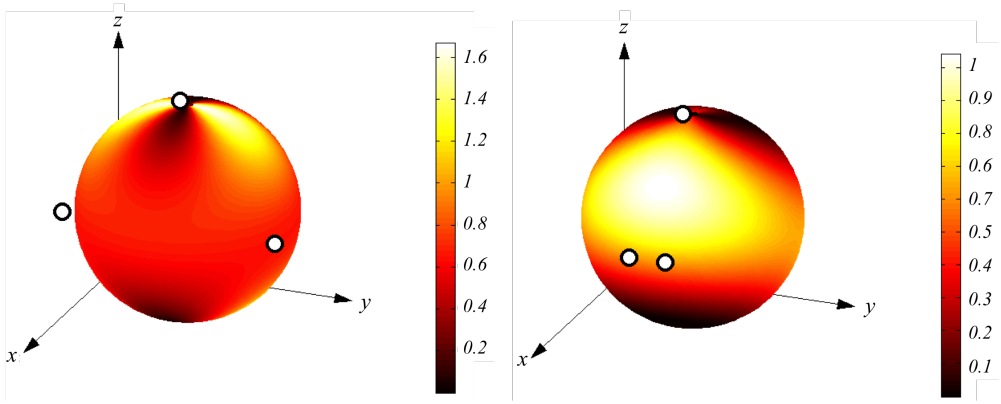
An observation of the variations of ρ^2 with respect to array configuration $\mathbf{\Omega}$ shows that this metric is related to the orthogonality of the considered source DOA regarding the array plane. As seen in Figure 5.4, the nulls of $(1-\rho^2)$ are located on DOAs coplanar to $M_1M_2M_3$, while it approaches 1 for values of $\mathbf{\Theta}$ orthogonal to $M_1M_2M_3$. This result is compliant with the observations on the CRB of the Uniform Linear Array discussed in Section 4.2.

r_θ^2 and r_φ^2 are to be considered as cost metrics, binding the considered array configuration to its "would have been configuration for a 1D DOA estimation for θ or φ estimations". These metrics define the optimal 1D DOA estimation efficiency once projected on the azimuthal or elevational planes. To be more specific, it is the whole term $||\mathbf{v}_\xi||^2(1-r_\xi^2)$ that verifies this proposition. The analysis is simpler for $\xi = \theta$ as the projection on the azimuthal plane is similar to the projection on Oxy while the elevational plane depends on θ. Figure 5.5 shows the variations of $||\mathbf{v}_\theta||^2(1-r_\theta^2)$ for two different configurations $\mathbf{\Omega} = (90^\circ, 20^\circ)$ and $(10^\circ, 20^\circ)$, presenting that the array optimum sensitivity shifts from $\theta = 0^\circ$ to $\theta = 90^\circ$ as β drops from 90° to 0°. The



(a) $\mathbf{\Omega} = (66^\circ, 12^\circ)$ (b) $\mathbf{\Omega} = (9^\circ, 12^\circ)$

Figure 5.5: Variation of $||\mathbf{v}_\theta||^2(1-r_\theta^2)$ on a sphere spanning (θ, φ) for configurations (a) $\mathbf{\Omega} = (66^\circ, 12^\circ)$ and (b) $(9^\circ, 12^\circ)$. The shift in sensitivity occurs when the projection of $\{M_1, M_2, M_3\}$ on the azimuthal plane (i.e. on Oxy) forms an equilateral triangle. See videos in //assets/partI/semiProdTheta/ for dynamic illustration.



(a) $\mathbf{\Omega} = (66^\circ, 12^\circ)$ (b) $\mathbf{\Omega} = (9^\circ, 12^\circ)$

Figure 5.6: Variation of $||\mathbf{v}_\varphi||^2(1 - r_\varphi^2)$ on a sphere spanning (θ, φ) for (a) $\mathbf{\Omega} = (66^\circ, 12^\circ)$ and (b) $(9^\circ, 12^\circ)$. The main lobe inversion occurs when the projection of $\{M_1, M_2, M_3\}$ on the elevational plane forms an equilateral triangle. See videos in //assets/partI/semiProdPhi/ for dynamic illustration.

shift occurs when the projection of $\{M_1, M_2, M_3\}$ in the azimuthal plane is an equilateral triangle, i.e. when $\mathbf{M_1M_2} \cdot \mathbf{Ox} = ||\mathbf{M_1M_2}||cos(\frac{\pi}{6})$. This configuration provides an isotropic sensitivity on θ estimations, i.e. the yellow/white area in Figure 5.5 circles the sphere, coplanar to Oxy. Note the drop of $||\mathbf{v}_\theta||^2(1-r_\theta^2)$'s maxima as the characteristic length of the array shifts from approx. half to a quarter wavelength (from approx. 1.6 to 0.4).

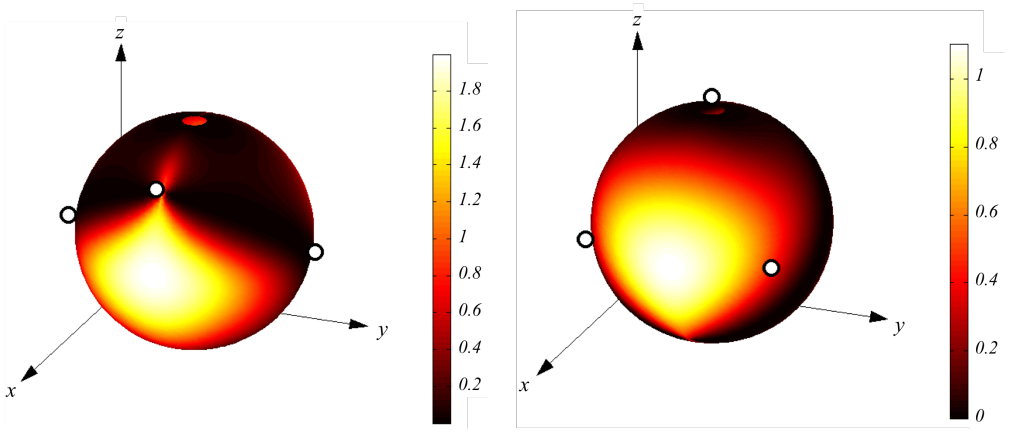
The same analysis holds for $\xi = \varphi$. As the projection of $\{M_1, M_2, M_3\}$ on the elevational plane is less intuitive, Figure 5.6 mimics Figure 5.5 for the term $||\mathbf{v}_\varphi||^2(1 - r_\varphi^2)$. As for θ, the general orientation of the antenna array toward Oyz or Ozx defines the maxima's position, partially conditioning φ estimation's accuracy.

These results are in accordance with those of Gazzah [76] and Dogandzic [65], who observed that the least and most desirable DOAs are perpendicular to each other and correspond to the antenna array main axes of symmetry, termed "*principal axes of inertia*" in [65].

Figures 5.7 and 5.8 show the variations of $||\mathbf{v}_\xi||^2(1 - r_\xi^2)(1 - \rho^2)$ as a function of (θ, φ), combining the effects of the two metrics. Except for the coefficient γ_M, these expressions are the inverse of the $\text{CRB}_{\xi,\xi}$ in (5.6), immediately related to the optimal accuracy of configuration $\mathbf{\Omega}$ regarding estimations on DOA $\mathbf{\Theta}$. γ_M has been discarded from the analysis as it does not depend on DOA.

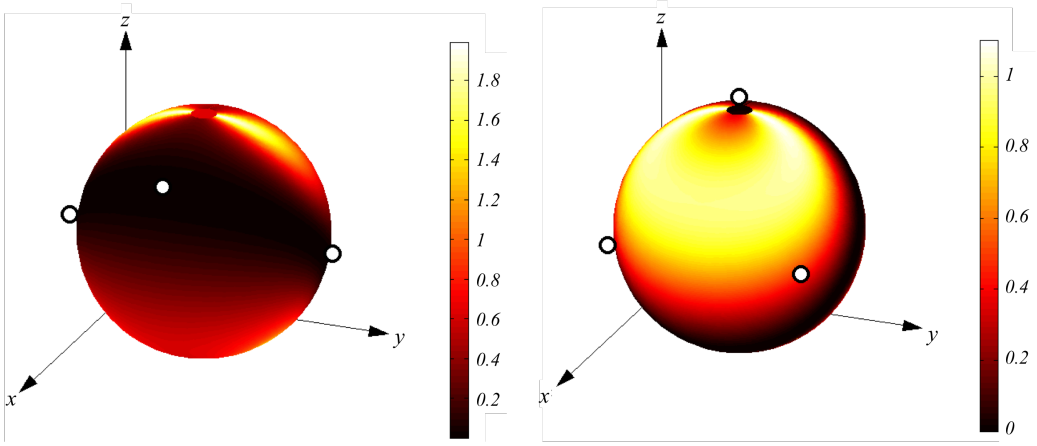
As ρ^2 and r_ξ^2 reflect the $\text{CRB}_{\xi,\xi}$ behavior for any array configuration $\mathbf{\Omega}$, they can be used to adapt any spherical array configuration to various search scenarios. For large distances, the victim's elevation with respect to antenna array approaches zero, while it may obviously be positioned anywhere in the horizontal plane. With no a priori knowledge of θ, the most favorable configuration is $\mathbf{\Omega} = \{120°, 90°\}$ to provide an isotropic precision on azimuthal estimations. This configuration is equivalent to a 3-element Uniform Circular Array (UCA). For building-like environments, $\mathbf{\Theta}$ is likely to assume values in the upper or lower front quadrants while the rescuer walks towards a victim potentially not on the same floor. The minima of $1/\text{CRB}_{\theta,\theta}$ can be iteratively oriented towards initially estimated DOAs (downwards or upwards as seen in Figures 5.7a and b) while maximizing $1/\text{CRB}_{\varphi,\varphi}$ values in the related quadrant (Figures 5.8a and b).

The next section completes this theoretical evaluation with simulations based on WINNER II and IlmProp propagation models. The objective is to assess the performance of the helmet-mounted geometry for a fixed $\mathbf{\Omega}$ configuration when faced with multi-path propagation, to emphasize (1) the difference between accuracy and robustness and (2) the share of the DOA estimation technique in the general performance of the DOA estimator.



(a) $\mathbf{\Omega} = (87^\circ, 51^\circ)$ (b) $\mathbf{\Omega} = (45^\circ, 0^\circ)$

Figure 5.7: Variation of $(1-\rho^2)||\mathbf{v}_\theta||^2(1-r_\theta^2)$ for (a) $\mathbf{\Omega} = (87^\circ, 51^\circ)$ and (b) $(45^\circ, 0^\circ)$. The general array orientation favors θ estimations on larger areas for DOAs orthogonal to the array plane. See videos in //assets/partI/invCRBTheta/ for dynamic illustration.



(a) $\mathbf{\Omega} = (87^\circ, 51^\circ)$ (b) $\mathbf{\Omega} = (45^\circ, 0^\circ)$

Figure 5.8: Variation of $(1-\rho^2)||\mathbf{v}_\varphi||^2(1-r_\varphi^2)$ for (a) $\mathbf{\Omega} = (87^\circ, 51^\circ)$ and (b) $(45^\circ, 0^\circ)$. The general array orientation favors φ estimations on larger areas for DOAs orthogonal to the array plane. See videos in //assets/partI/invCRBPhi/ for dynamic illustration.

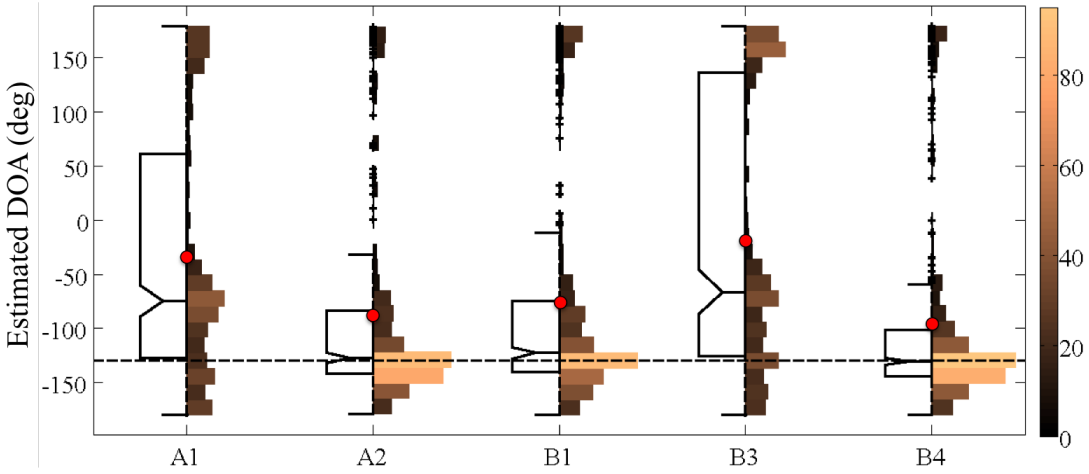


Figure 5.9: Results of a Monte Carlo simulation on DOA estimations for various scenarios of the WINNER II model (see Table 4.2). The DOA estimation is based on MUSIC coupled with a 3-elements UCA array, with each boxplot presenting results for 200 estimations. Black dotted line represents the targeted DOA to estimate, red dots the means of the estimated DOAs. Histograms illustrate the exact repartition of DOA estimates. The most erratic behaviors are observed for scenarios A1 and B3 (indoor office and large indoor hall).

5.4 Evaluation Based on Propagation Models

Let us assume a scenario where the emitter is positioned on the same horizontal plane as the antenna array. Based on the results of the previous section, the optimal array configuration for an unknown azimuth is $\mathbf{\Omega} = (120°, 90°)$. This geometry is equivalent to a 3-element UCA and provides an isotropic precision for horizontal DOA estimations. For the remaining of this section, this geometry is coupled with a MUSIC based DOA estimation technique.

Figure 5.9 shows the average performance of this DOA estimator for various WINNER II scenarios in noiseless conditions. While it produces acceptable performance in "outdoor" environments, it is not robust to dense multi-path propagation specific to indoor environments. Not shown here, the estimator's accuracy further decreases as the SNR drops.

An assessment based on the IlmProp raytracing model produces similar results. Two scenarios have been designed, the first simulating a simple urban environment, the second a crowded indoor environment. Prior to the assessment, a grid of potential positions for the receiver (DOA estimator) is defined for both environment, as explained in Section 4.3.3. The impulse responses (IRs) from emitter to receiver potential position on grid nodes are simulated and stored. After defining the initial position of the

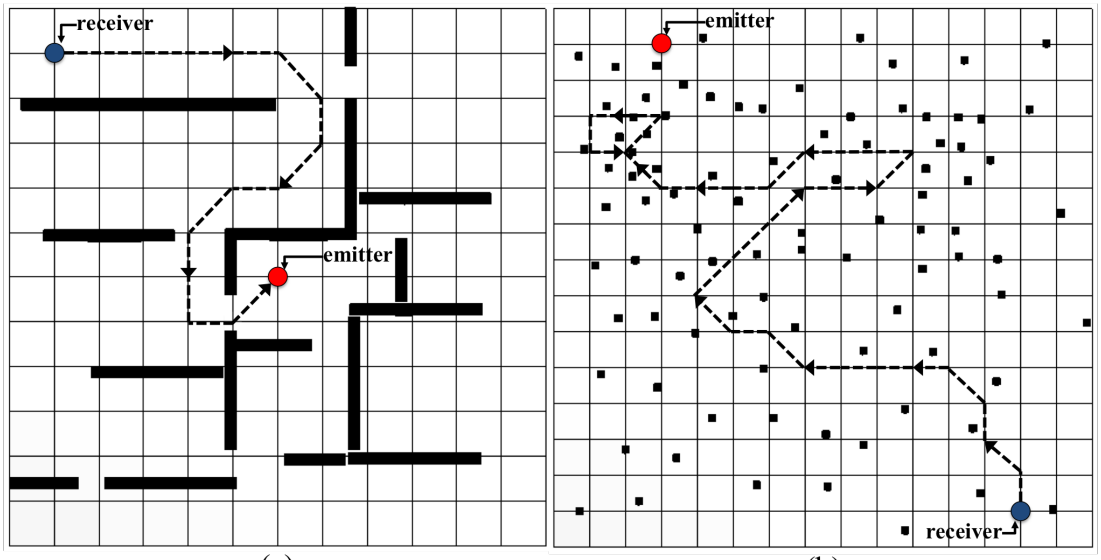


Figure 5.10: Automatic navigation based on DOA estimator output for two propagation scenarios simulated in IlmProp. The DOA estimation is based on MUSIC coupled with a 3-elements UCA array. The automated navigation simply consists in moving the receiver (DOA estimator) to the neighboring grid node that best fits the estimated DOA. (a) Urban scenario. (b) Indoor environment, each black square is a pillar designed to reflect RF waves, designed to simulate a dense multi-path environment. In the second scenario, the automated navigation is stuck in a loop due to the impact of multi-path propagation on MUSIC estimations.

receiver on the grid, the evaluation consists in iteratively moving the receiver to the neighboring grid node based on its DOA estimation issued from the current node's IR. As expected, the DOA estimator easily reaches the emitter as long as the environment does not present dense multi-path propagation, see Figure 5.10.

Both simulations indicate that the considered DOA estimator becomes erratic in dense multi-path environments. Arranging the configuration of the array elements as suggested in the previous sections does not improve these results, even when one uses the true DOA to optimize antennas' position. To use the helmet-mounted geometry of Figure 5.1 in dense multi-path propagation conditions, the only available option is to work on the DOA estimation technique to improve the robustness of the DOA estimator.

As presented in Chapter 3, some algorithms exist to make an estimator more robust to the presence of correlated signals. The most notable are the "forward-backward averaging" [78] and the "spacial smoothing" [79], two preprocessing schemes that take advantage of specific properties of the array geometry to differentiate correlated sig-

nals prior to the estimation. UCA/spherical geometries are not compatible with the basic versions of these algorithm, as they rely on the centro-hermitian property of the beamspace array manifold (or equivalently its estimated spatial covariance matrix) [80]. Some studies have adapted forward-backward averaging [52, 53] and spatial smoothing [54] to UCA geometries to reduce the impact of coherent signals on estimation accuracy. However, these techniques require geometries with a large number of antenna elements, much more than the three antennas proposed for the currently considered array.

5.5 Conclusion

This chapter presented a case-study evaluation and optimization of an antenna array configuration. Designed for DOA estimation during search and rescue applications, the design proposed consists in mounting a lightweight antenna array on a rescuer's helmet to optimize RF reception quality.

An intuitive formulation of the FIM specific to spherical geometries was proposed, from which were extracted the CRBs related to azimuth and elevation estimations. The CRBs were decomposed into metrics, used to understand the impact of antennas positioning on DOA estimations accuracy. This approach allowed for the optimization of antenna positioning through basic geometrical considerations, adapting the array configuration to specific DOA estimation scenarios.

The antenna array was coupled with a MUSIC estimation technique to assess its performance in realistic propagation conditions. Simulations were conducted in WINNER II and IlmProp, two different propagation models introduced in Chapter 4. While the considered DOA estimator produced acceptable results in outdoor environments, it lacked robustness when faced with typical indoor multi-path propagation. Its improvement would require the implementation of more advanced DOA estimation techniques, limited by the geometry and the number of antennas of the considered array. The modification of the array itself is eventually limited by the targeted lightweight application.

5.6 Appendix: Derivation of the FIM

This appendix presents an exhaustive demonstration of the FIM formulation of (5.3) used to establish $\mathrm{CRB}_{\xi,\xi}$ scalar expressions in (5.6).

According to the hypothesis stated in Section 5.2, the probability density functions (pdf) $f_{\mathbf{x}}$ of received snapshots $\mathbf{x}[t]$ along the N measurements are independent. Their joint pdf thus takes the form

$$\prod_{t=1}^{N} f_{\mathbf{x}}(\mathbf{x}[t]|\mathbf{\Theta}, \sigma_s^2, \sigma_n^2) = \frac{1}{\pi^{MN}} \frac{1}{|\mathbf{R}_x|^N} exp(-\sum_{t=1}^{N} \mathbf{x}^H[t]\mathbf{R}_x^{-1}\mathbf{x}[t]) \tag{5.7}$$

where the assumptions made on $s[t]$ and $\mathbf{n}[t]$ imply, using the steering vector simplified notation $\mathbf{a} = \mathbf{a}_{\mathbf{\Omega}}(\mathbf{\Theta})$:

$$\mathbf{R}_x = \sigma_s^2 \mathbf{a}\mathbf{a}^H + \sigma_n^2 \mathbf{I}_N \tag{5.8}$$

$\mathbf{R}_x$ being the *spatial* covariance matrix of vector $\mathbf{x}$. Since the exponential term in (5.7) is a scalar, we can use the circular permutation invariance of the trace operator to reveal the covariance estimation matrix $\hat{\mathbf{R}}_x$:

$$\begin{aligned} \sum_{t=1}^{N} \mathbf{x}^H[t]\mathbf{R}_x^{-1}\mathbf{x}[t] &= \mathrm{tr}(\sum_{t=1}^{N} \mathbf{x}[t]\mathbf{x}^H[t]\mathbf{R}_x^{-1}) \\ &= \mathrm{tr}(N\ \hat{\mathbf{R}}_x\ \mathbf{R}_x^{-1}) \end{aligned} \tag{5.9}$$

where

$$\hat{\mathbf{R}}_x = \frac{1}{N} \sum_{t=1}^{N} \mathbf{x}[t]\mathbf{x}[t]^H \quad \text{with} \quad \mathrm{E}\left\{\hat{\mathbf{R}}_x\right\} = \mathbf{R}_x$$

The basic formulation of the log-likelihood function is then obtained by replacing (5.9) in (5.7) and taking its natural logarithm

$$\Lambda = -MN\log(\pi) - N\log|\mathbf{R}_x| - N\mathrm{tr}(\mathbf{R}_x^{-1}\hat{\mathbf{R}}_x) \tag{5.10}$$

Λ thus needs to be shaped for derivation regarding $\mathbf{\Theta}$. Using (5.8) to replace $|\mathbf{R}_x|$

$$\begin{aligned} |\mathbf{R}_x| = |\sigma_s^2 \mathbf{a}\mathbf{a}^H + \sigma_n^2 \mathbf{I}_M| &= \sigma_n^{2(M-1)}(\sigma_s^2 \mathbf{a}\mathbf{a}^H + \sigma_n^2) \\ &= \sigma_n^{2M}(\alpha M + 1)1 \end{aligned} \tag{5.11}$$

where $\alpha = \sigma_s^2/\sigma_n^2$. The right hand term of (5.10) can also be simplified by applying the Woodbury inversion formula on $\mathbf{R}_x^{-1}$

$$\begin{aligned} \mathrm{tr}(\mathbf{R}_x^{-1}\ \hat{\mathbf{R}}_x) &= \frac{1}{\sigma_n^2}\mathrm{tr}([\alpha\mathbf{a}\mathbf{a}^H + \mathbf{I}_M]^{-1}\hat{\mathbf{R}}_x) \\ &= \frac{1}{\sigma_n^2}\mathrm{tr}([\mathbf{I}_M - \frac{\alpha\mathbf{a}\mathbf{a}^H}{1+\alpha M}]\ \hat{\mathbf{R}}_x)1 \end{aligned} \tag{5.12}$$

Injecting both (5.11) and (5.12) in (5.10), the log-likelihood becomes

$$\begin{aligned}\Lambda = &-MN(\log(\pi)+\log(\sigma_n^2)) - \frac{N}{\sigma_n^2}\mathrm{tr}(\hat{\mathbf{R}}_x) \\ &-N\log(1+\alpha N) + \frac{N}{\sigma_n^2}\frac{\alpha}{(1+\alpha M)}\mathbf{a}^H\hat{\mathbf{R}}_x\mathbf{a}\end{aligned} \tag{5.13}$$

Since the considered derivation concerns $\boldsymbol{\Theta}$ dependent elements, only the last term of (5.13) is of interest. For the sake of clarity, derivatives such as $\delta_\xi(\mathbf{a})$ and $\delta^2_{\xi,\zeta}(\mathbf{a}^H)$ will thereafter be denoted as $\delta_\xi\mathbf{a}$ and $\delta^2_{\xi,\zeta}\mathbf{a}^H$, the derivative affecting its adjacent term only. The first derivative of (5.13) with respect to $\xi \in \{\theta, \varphi\}$ is then

$$\delta_\xi\Lambda = \frac{N}{\sigma_n^2}\frac{\alpha}{(1+\alpha M)}(\delta_\xi\mathbf{a}^H\hat{\mathbf{R}}_x\mathbf{a} + \mathbf{a}^H\hat{\mathbf{R}}_x\delta_\xi\mathbf{a})$$

the second derivative related to $\zeta \in \{\theta, \varphi\}$ naturally reveals cross conjugated terms, hereafter gathered in $\Re\{\cdot\}$

$$\delta^2_{\xi,\zeta}\Lambda = 2\frac{N}{\sigma_n^2}\frac{\alpha}{(1+\alpha M)}(\Re\{\delta^2_{\xi,\zeta}\mathbf{a}^H\hat{\mathbf{R}}_x\mathbf{a}\} + \Re\{\delta_\xi\mathbf{a}^H\hat{\mathbf{R}}_x\delta_\zeta\mathbf{a}\}) \tag{5.14}$$

According to the FIM definition stated in Section 5.3.1, we need to evaluate the negative expectation of the second order derivative log-likelihood. Replacing the estimate $\hat{\mathbf{R}}_x$ by its true value $\mathbf{R}_x$ in (5.14) gives

$$\mathrm{E}\left\{-\delta^2_{\xi,\zeta}\Lambda\right\} = -2\frac{N}{\sigma_n^2}\frac{\alpha}{(1+\alpha M)}(\Re\{\delta^2_{\xi,\zeta}\mathbf{a}^H\mathbf{R}_x\mathbf{a}\} + \Re\{\delta_\xi\mathbf{a}^H\mathbf{R}_x\delta_\zeta\mathbf{a}\})$$

which once developed using $\mathbf{R}_x$ expression in (5.8) gives

$$\mathrm{E}\left\{-\delta^2_{\xi,\zeta}\Lambda\right\} = -\frac{2\alpha N}{(1+\alpha M)}\left[(1+\alpha M)\Re\{\delta^2_{\xi,\zeta}\mathbf{a}^H\mathbf{a}\} + \Re\{\delta_\xi\mathbf{a}^H\delta_\zeta\mathbf{a}\} + \alpha\Re\{\delta_\xi\mathbf{a}^H\mathbf{a}\mathbf{a}^H\delta_\zeta\mathbf{a}\}\right] \tag{5.15}$$

Let us then develop the differential terms $\delta_\xi\mathbf{a}^H$ and $\delta^2_{\xi,\zeta}\mathbf{a}^H$, introducing the notations

$$\mathbf{a} = \begin{pmatrix} e^{j\frac{2\pi r}{\lambda}u_1} \\ e^{j\frac{2\pi r}{\lambda}u_2} \\ \vdots \\ e^{j\frac{2\pi r}{\lambda}u_M} \end{pmatrix} \quad \text{and} \quad \mathbf{U} = \begin{pmatrix} u_1 & 0 & \dots & 0 \\ 0 & u_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_M \end{pmatrix}$$

resulting in the first and second order derivatives:

$$\begin{aligned}\delta_\xi\mathbf{a}^H &= -j\frac{2\pi r}{\lambda}\mathbf{a}^H\delta_\xi\mathbf{U} \\ \delta^2_{\xi,\zeta}\mathbf{a}^H &= -j\frac{2\pi r}{\lambda}\mathbf{a}^H\delta^2_{\xi,\zeta}\mathbf{U} - (\frac{2\pi r}{\lambda})^2\mathbf{a}^H\delta_\xi\mathbf{U}\delta_\zeta\mathbf{U}\end{aligned}$$

used to develop (5.15) into

$$-E\left\{\delta^2_{\xi,\zeta}\Lambda\right\} = \frac{\gamma_M}{M}(\mathbf{a}^H\mathbf{a}\mathbf{a}^H\delta_\zeta\mathbf{U}\delta_\xi\mathbf{U}\mathbf{a} - \mathbf{a}^H\delta_\zeta\mathbf{U}\mathbf{a}\mathbf{a}^H\delta_\xi\mathbf{U}\mathbf{a}) \tag{5.16}$$

with

$$\gamma_M = 2\alpha T\frac{1}{\left(1+\frac{1}{\alpha M}\right)}\left(\frac{2\pi r}{\lambda}\right)^2$$

It is noted here that $E\left\{\delta^2_{\xi,\zeta}\Lambda\right\} = E\left\{\delta^2_{\zeta,\xi}\Lambda\right\}$ since $\mathbf{U}$ is diagonal with real coefficients and $\mathbf{a}\mathbf{a}^H$ is a hermitian matrix. By noticing the weighted scalar product in (5.16):

$$\begin{aligned}(\mathbf{a}^H\mathbf{a}\mathbf{a}^H\delta_\zeta\mathbf{U}\delta_\xi\mathbf{U}\mathbf{a} - \mathbf{a}^H\delta_\zeta\mathbf{U}\mathbf{a}\mathbf{a}^H\delta_\xi\mathbf{U}\mathbf{a}) &= \mathbf{a}^H\mathbf{a}(\mathbf{v}_\zeta^H\mathbf{\Pi}_{\perp\mathbf{a}}\mathbf{v}_\xi)\\ &= M\langle\mathbf{v}_\xi,\mathbf{v}_\zeta\rangle_{\perp\mathbf{a}}\end{aligned}$$

where $\mathbf{\Pi}_{\perp\mathbf{a}}$ is the orthonormal projector on the *noise space* defined in the MUSIC subspace decomposition

$$\mathbf{\Pi}_{\perp\mathbf{a}} = \mathbf{I}_\mathbf{M} - \tilde{\mathbf{a}}\tilde{\mathbf{a}}^H \quad \text{with} \quad \tilde{\mathbf{a}} = \frac{\mathbf{a}}{\sqrt{\mathbf{a}^H\mathbf{a}}}$$

the right-hand term in (5.16) becomes

$$-E\left\{\delta^2_{\xi,\zeta}\Lambda\right\} = \gamma_M\langle\mathbf{v}_\xi,\mathbf{v}_\zeta\rangle_{\perp\mathbf{a}} \tag{5.17}$$

which proves the proposed FIM in (5.3):

$$F(\mathbf{\Theta}) = \gamma_M\begin{pmatrix}\langle\mathbf{v}_\theta,\mathbf{v}_\theta\rangle_{\perp\mathbf{a}} & \langle\mathbf{v}_\theta,\mathbf{v}_\varphi\rangle_{\perp\mathbf{a}}\\ \langle\mathbf{v}_\varphi,\mathbf{v}_\theta\rangle_{\perp\mathbf{a}} & \langle\mathbf{v}_\varphi,\mathbf{v}_\varphi\rangle_{\perp\mathbf{a}}\end{pmatrix}$$

or its equivalent vectorial form that highlights the MUSIC noise space projection

$$F(\mathbf{\Theta}) = \gamma_M\begin{pmatrix}\mathbf{v}_\theta^H\\ \mathbf{v}_\varphi^H\end{pmatrix}\mathbf{\Pi}_{\perp\mathbf{a}}\begin{pmatrix}\mathbf{v}_\theta & \mathbf{v}_\varphi\end{pmatrix}$$

Hence, a simple inversion of $F(\mathbf{\Theta})$ reveals the determinant

$$(1-\rho^2) = \langle\mathbf{v}_\theta,\mathbf{v}_\theta\rangle_{\perp\mathbf{a}}\langle\mathbf{v}_\varphi,\mathbf{v}_\varphi\rangle_{\perp\mathbf{a}} - \langle\mathbf{v}_\theta,\mathbf{v}_\varphi\rangle_{\perp\mathbf{a}}\langle\mathbf{v}_\varphi,\mathbf{v}_\theta\rangle_{\perp\mathbf{a}}$$

and leads to the CRB matrix, which holds both $\mathrm{CRB}_{\xi,\xi}$ introduced in (5.6) in its upper left and bottom right diagonal terms

$$F(\mathbf{\Theta})^{-1} = \frac{1}{\gamma_M}\frac{1}{(1-\rho^2)}\begin{pmatrix}\langle\mathbf{v}_\theta,\mathbf{v}_\theta\rangle^{-1}_{\perp\mathbf{a}} & -\rho^2\langle\mathbf{v}_\theta,\mathbf{v}_\varphi\rangle^{-1}_{\perp\mathbf{a}}\\ -\rho^2\langle\mathbf{v}_\varphi,\mathbf{v}_\theta\rangle^{-1}_{\perp\mathbf{a}} & \langle\mathbf{v}_\varphi,\mathbf{v}_\varphi\rangle^{-1}_{\perp\mathbf{a}}\end{pmatrix}$$

where $\langle \mathbf{v}_\xi, \mathbf{v}_\xi \rangle_{\perp \mathbf{a}} = \mathbf{v}_\xi^H (\mathbf{I_M} - \tilde{\mathbf{a}}\tilde{\mathbf{a}}^H)\mathbf{v}_\xi = ||\mathbf{v}_\xi||(1 - r_\xi^2)$ for $\xi \in \{\theta, \varphi\}$.

Chapter 6

Conclusion on DOA Estimation

Part I of this manuscript has outlined the fundamentals of Direction Of Arrival (DOA) estimation: the main estimation techniques, the rules behind the design of an antenna array, and how to evaluate the performance of a DOA estimator. The case-study presented in Chapter 5 employed the techniques presented in Chapter 3 to design a lightweight DOA estimator for search and rescue applications. The tools introduced in Chapter 4 were employed to optimize its geometry and assess its performance in various propagation scenarios.

The optimization criterion was based on the Cramer-Rao Bounds associated to the antenna array, modifying the array geometry to concentrate the array's precision with respect to specific DOAs. Despite this optimization, the DOA estimator obtained by coupling the antenna array to the high-resolution MUSIC estimation technique exhibited a poor degree of robustness regarding DOA estimation in dense multi-path environments. The case-study concluded that there was a need for advanced high-resolution techniques (other than MUSIC) or classical smoothing methods [78, 79] to improve the robustness of the estimation regarding multi-path propagation, however their use was limited by the reduced number of antennas in the considered array.

The limitation imposed by the number of antennas in the array on DOA estimation techniques is a known issue in the design of portable DOA estimators [81]. The use of classical signal processing techniques to improve the robustness of the DOA estimator to multi-path propagation generally requires arrays composed of a minimum of eight to twelve antennas [55]. As discussed in Chapter 3, the number of antennas more or less defines the final size of the array. Both spatial diversity and aliasing must be considered in order to optimize the contribution of each antenna in the DOA estimation, which often results in inter-element spacing values close to half the wavelength of the received signal. For DOA estimation of GSM signals where $\lambda/2 \approx 16$ cm, an array with more than four or five antennas is likely to hinder rescuer's movements during their search because of its size[1].

Recent studies have reported promising results for DOA estimation on correlated signals based on low-profile arrays. The work reported in [82] showed that a 36-elements

[1] for GSM-900 frequencies, see Chapter 2

spherical design of radius $\lambda/3$ coupled with a modified version of the ESPRIT algorithm could produce robust and accurate estimations when faced with multi-path propagation. The most accomplished DF design published to date regarding search and rescue operation is perhaps the 50×50 cm^2 5-element T-shaped array proposed in [81]. Combined with a DOA estimation technique based on an improved form of MUSIC [17], this design proved to be robust to multi-path propagation conditions and allowed users to find victims distributed throughout a 100×100 m^2 area in a few minutes with a resolution of 2 to 5 meters.

These studies suggest that it is possible to design a reasonably low-profile DOA estimator for search and rescue operations, while further research would be required to produce robust estimations with a helmet or hand sized array, compact enough to be effortlessly carried along during GSM-based search and rescue operations. The reminder of this thesis will assume that such a DOA estimator can be implemented to investigate the impact of a sonification on the performance of the final Direction Finder (DF) as a navigation aid. As discussed in Chapter 1, the objective of this research is to design a DF whose performance relies as much on user's listening abilities as on raw computer-based DOA estimation efficiency. For that purpose, Parts II and III will not only consider DF designs based on automated DOA estimation but will also consider the sonification and prototyping of simpler DF designs, e.g. involving manual DOA estimations based on the signal received by a unique a directional antenna, where they to prove to be more efficient than complex DOA estimators when coupled with an intuitive sonification paradigm.

If one was to focus on the optimization of the DOA estimator design suggested in Chapter 5, future work would first involve the adaptation of the DOA estimation technique proposed in [81] to the suggested helmet-mounted array geometry. After performance validation in realistic propagation conditions, the impact of the helmet/head on radio frequency propagation would have to be assessed as the non-isotropic difference in propagation path between antennas is likely to alter the behavior of the DOA estimator. Finally, the impact of slight differences between modeled and real array geometries regarding the accuracy of DOA estimation would have to be assessed, given the demanding and hazardous conditions of the targeted application[2].

[2] see Section 3.4.

Part II

Sonification

The work presented here concerns research carried out related to sonification of information issued from the direction finder. An overview of the related research field is provided along with an introduction to the essential notions and techniques employed in sonification. Various sonification paradigms are introduced and discussed as potential direction finder output, from simple received signal strength to estimated direction of arrival. The last chapter in this part presents a study on reducing the impact of noise in real-time data exploration based on psychoacoustics applied to parameter mapping sonification.

Chapter 7

The Use of Sound to Convey Information

Contents

Some would evoke survival mechanisms, other would describe its relationship to emotions or its impact on personal development, the fact is that humans continuously use sound to gather information on their surroundings. From there, its only natural to assume that sound has a role to play in modern interfaces to convey information.

This chapter presents an overview of the different techniques and problematics related to the display of information through auditory streams, a process referred to as *sonification*. Section 7.1 presents fundamental notions of sound perception and introduces the most common sonification techniques. Section 7.2 contains a literature review, focused on research and applications similar to the task of the current project. As the objective is to design an audio direction finder, which can roughly be described as "an electromagnetic probe with an audio interface", said task can be reduced to the **interactive** sonification of a **real-time stream** of **data**. For a more exhaustive survey on sound perception, sonification, and psychoacoustics, see [21, 83–86]. The review proposed in [87] provides a solid introduction on the design and evaluation of auditory interfaces.

7.1 Fundamentals of Sonification

Several basic features of auditory perception suggest that sound can be effective at representing data, e.g. compared to vision [83]. Visualization is often taken as a reference to weight the pros and cons of sonification, as it is the most commonly employed modality for displaying information.

First of all, the perception of a sound does not require one to face its source, nor does the evaluation of its position in space. Human hearing is also well suited to spot subtle changes in complex repeated sound patterns or to simultaneously monitor and process multiple auditory streams [88]. The use of sound in surgical blocks is often taken as a case study for sonification designs, as it requires a constant evaluation of patient conditions from various machines while focusing on the operating table [89,90]. Sonification design concerns the encoding or "mapping" of information onto sound, keeping these features in mind to make the most of our listening capabilities.

7.1.1 Sound Parameters

Much like color, shape, or placement in visual displays, sound designers will basically work with the following sound parameters:

Pitch
Pitch is the perceptual equivalent of frequency, generally related to the relative position of a tone on a musical scale. Where frequency is an objective, scientific concept, pitch is a subjective psychoacoustic attribute of sound. As such, the perceived pitch of a sound depends on its intensity, duration, and on other interfering sounds [91,92].

Pitch is sometimes expressed in "mel", referring to a perceptual scale proposed by Stevens et al. in [93]. To create this scale, a reference point between Hertz and mels was arbitrarily defined by assigning a perceptual pitch of 1000 mels to a 1 kHz pure and continuous tone at 40 dB SPL[1] above the listener's hearing threshold. An experiment was then designed where listeners had to listen to that reference tone and increase (or decrease) its frequency until they felt the pitch was twice as high (or as low) as the original sound. The resulting mel scale is illustrated in Figure 7.1a. Sonification designers often prefer pitch over frequency as it takes into account listeners' perception.

Loudness
Loudness is the perceptual equivalent of sound intensity (i.e. sound pressure level). The perceived loudness depends on pitch, sound duration, predictability, and fatigue [84].

Stevens et al. [94] followed the exact same methodology than for pitch to define a

[1] Sound Pressure Level (SPL) is a unit related to the intensity of a sound, measured in dB, defined as $20\log_{10}(p_{\text{rms}}/p_0)$, where p_{rms} is the root mean square sound pressure, expressed in Pa and p_0 is the reference sound pressure, generally defined as 20 µPa in air.

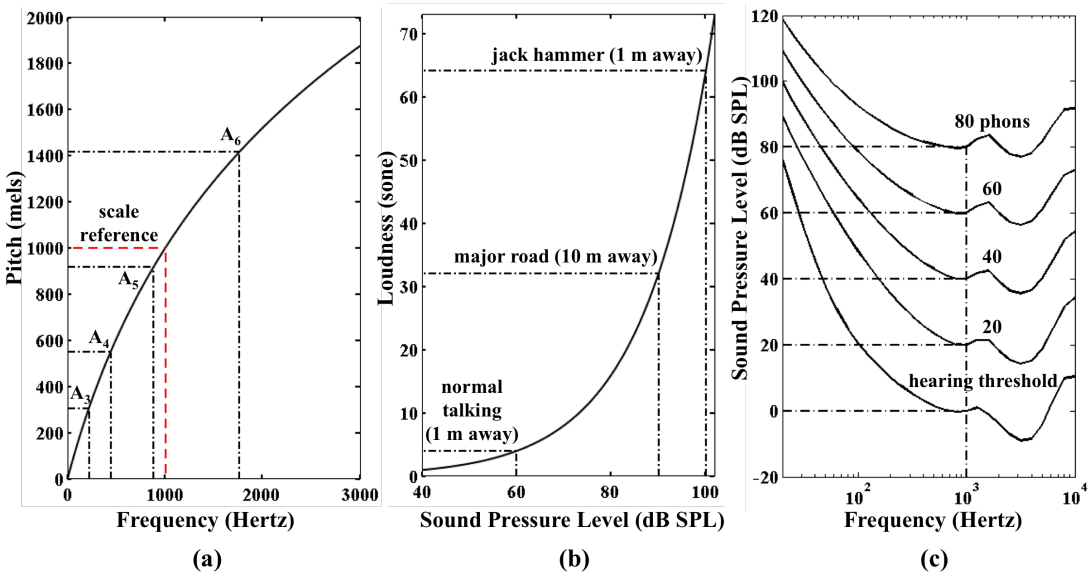


Figure 7.1: (a) mel as a function of frequency after the expression proposed in [95]: $f_{mel}(\text{freq}) = 2595 \log_{10}(1 + \text{freq}/700)$. The reference between both units is 1000 mels = 1 kHz. The values of notes A_i are indicated for reference, e.g. A_4 at 440 Hz. (b) sone as a function of sound pressure level after the approximation proposed in [84]: $f_{sone}(L_N) = 2^{\frac{L_N - 40}{10}}$ where L_N is expressed in "phons". This approximation holds for $L_N > 40$ dB SPL. By definition, the number of phon of a sound is the SPL in dB of a pure 1 kHz tone that sounds just as loud, see (c). The reference between both units is arbitrarily defined as 1 sone = 40 dB SPL. (c) Perceived loudness as a function of frequency after ISO 226 [96]. These "equal-loudness contours" allow the mapping of the SPL of a pure tone to its perceived loudness regardless of its frequency.

perceptual scale for perceived amplitude based on the "sone": a loudness of 1 sone was defined as the perceived intensity of a 1 kHz continuous pure tone at 40 dB SPL. As with pitch, loudness is preferred over amplitude in sonification design for its subjective aspect.

Tempo and Rhythm

Both notions refer to the temporal variations of a sound. Tempo represents the period of a sound event, generally expressed in "bpm" (beats per minute). Rhythm on the other hand defines the perceived temporal organization of a sound pattern, a subjective notion which eventually leads to complex considerations on affect and grouping [97,98].

Timbre

Timbre is related to the physical notions of frequency spectrum and temporal envelope

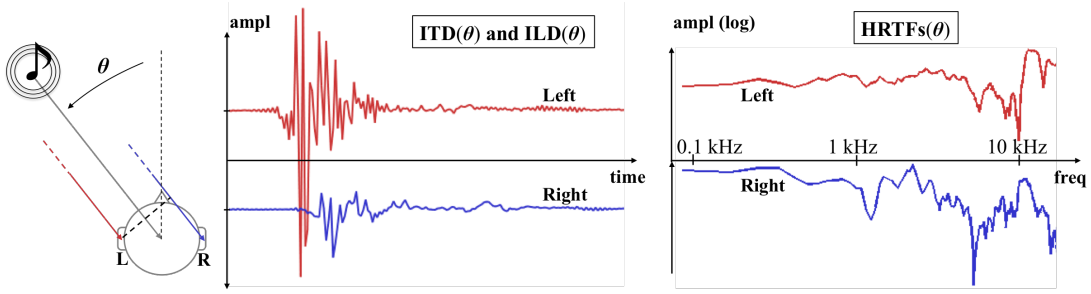


Figure 7.2: Illustration of the spatial cues induced by the difference in propagation path between left and right ears. For the sake of illustration, only the azimuth θ of the sound source is considered.

[99]. Often associated with subjective terms such as "coloration" or "warmth", timbre concerns all the things that make two instruments sound different while playing the same note at equal intensity.

Spatial Location

The mechanisms involved in the perception of auditory events in space have been the subject of many studies concerning both the analysis and reproduction of what is referred to as *spatial hearing* [21, 100]. Our ability to localize sounds in space is understood to be issued from a learning process involving the analysis of auditory cues based on the feedback from other modalities (e.g. visual or haptic) [101, 102]. As illustrated in Figure 7.2, a sound coming from a given location propagates through different paths before reaching the left and right eardrums. This difference in propagation path induces an Interaural Time Difference (ITD) along with some frequency selective shadowing or resonance due to propagation medium, head, shoulders, pinna, etc. This location-specific or "spatial" frequency filtering is generally described through a set of filters called the Head-Related Transfer Function (HRTF) [21, 103]. Specific to left and right ears of each individual, an HRTF contains the spatial filtering information for any source point (usually distributed on a sphere) to the ear canal. Before the impact of this frequency-selective filtering on spatial hearing was understood, only the average difference of sound pressure level between left and right ears was considered, defined as the Interaural Level Difference (ILD). Compared to the ITD that varies only as a function of the azimuth and elevation of a sound source, the ILD and the HRTF also depends on the distance between the source and the listener, as the difference in SPL between the two ears varies with the distance already travelled by the acoustic wave in the propagation medium.

Studies have shown that humans selectively use HRTF and ITD related cues to localize sound sources depending on their frequency components [21]. For frequencies below

about 1 kHz, localization is based on ITD and does not yield good results for elevation estimations. Sources composed of frequency higher than 3 kHz are usually localized much more precisely, based on HRTF related cues induced by the frequency filtering of the head and pinnae. These two different localization mechanisms can be explained by considering the difference in sound wavelength compared to ITD and HRTF "*characteristic length*" [104]. The best localization results are usually obtained for sound sources that present a broad spectrum and a short attack time, typically repeated bursts of white noise. For sources with similar frequency and time parameters, sounds which the listeners are used to hearing (e.g. human voice) will generally be more accurately localized.

The concepts of HRTF and ITD are mainly employed for the reproduction of spatial hearing via headphones, technique known as *binaural synthesis*. The main issues of this technique are related to the inconsistencies between the production model and the reality that weaken the perception of spatial sound:

- Non-individual HRTF: because the brain is adapted to a specific listener's morphology (shoulders, head, pinnae, etc.) the use of non-individual HRTFs provokes sound internalization, front-back confusions, and confused elevation sensations [105, 106].
- Non-tracked head movements: the brain is used to processing sound localization in the presence of brief head movements. If the HRTFs are not accordingly updated the object will be perceived as attached to listener's head or internalized (felt in between left and right ear) [107].
- Room acoustics: to reduce computation cost, binaural synthesis often ignores the simulation of sound propagation in real environments (never mind the listener's actual environment). As the brain learnt to identify the spectrum coloration related to different environment configurations (early reflections, diffusion, etc.), sounds issued from synthesis that does not take into account room acoustics will often be perceived as *flat* or inconsistent [107].
- No cross-modal feedback: ultimately, the absence of any visual or haptic source will prevent the last step of the localization process, known as the "*ventriloquism*" or "*visual capture*" effect [108], when subjects associate the origin of a sound to a visual object. Studies on cross-modal localization indicated that visual information were prone to over-ride auditory cues, e.g. through experiments where participants had to localize a sound source while presented with a visual distractor [108]. The study presented in [109] particularly emphasizes the dependence of auditory localization on constant feedback from non-auditory cues, observing a surprising ability of participants to adjust their auditory perception to new HRTF after 10 to 60 days of chronic exposure.

7.1.2 Sonification Techniques

The choice of a sonification technique mainly depends on the available data and the nature of the information to convey. From simple event notification to complex data stream analysis, researchers in the field of auditory display have developed the following techniques to address various application scenarios:

Notification Sounds (Earcons, Auditory Icons and Spearcons)
These sonifications are usually designed as short sounds to inform the user about a system state. From message notifications to folders creation [110], earcons and auditory icons are the sounds one grows so familiar with through day to day interactions with Human Computer Interfaces (HCIs).

Introduced by Blattner in 1989 [111], earcons are defined as "*a brief succession of pitches arranged to produce a rhythmic and tonal pattern sufficiently distinct to function as an individual recognizable entity*". The concept itself is not new though, as it was already used in e.g. Morse code during 19^{th} century where the repetition pattern of a single pitched tone was used to encode a character sequence.

Auditory icons resemble earcons yet they imply a natural affordance between the sound and the event, such as the crumpling of a piece of paper to represent the suppression of a document from a computer [112]. As auditory icons are based on meaningful non-speech sounds, they typically require a shorter learning process when compared to earcons [113].

Finally, the term spearcon [114] refers to short sounds based on speed up speech. Their more successful application is probably the design of interfaces for visually impaired users [113]. Spearcons involve almost no learning process (but learning the related language and the ability to hear "quickly") for a nearly infinite description diversity.

Data Sonification (Audification, PMSon and MBSon)
Audification, Parameter Mapping Sonification (PMSon) and Model-Based Sonification (MBSon) are sonification techniques used for data exploration and monitoring.

Audification consists in a "*direct translation of a data waveform into sound*" [115]. Its applications range from EEG[2] sonification [116] to "Auditory Seismology", where seismograms are *audified* by accelerating their playback speed to shift their frequency in the human audio range [117]. Audification seldom involves any preprocessing other than time or frequency shift, dilatation, and filtering, relying entirely on the strengths of the auditory system to explore and analyze the data. This sonification technique is particularly efficient for identifying irregularities in large datasets with periodic components,

[2]Electro-Encephalogram: recording of electrical activity along the scalp.

replacing a tedious visual search by a mere listening of soundscape singularities [118].

PMSon consists in the mapping of data values onto sound parameters (e.g. pitch, tempo, spatial position, etc.). Modern cars parking aids (proximity sensors' sonification), Wall-E's charge control system or Apple's laptops volume control feedback are all basic illustrations of system's state monitoring through PMSon. PMSons of parking aids usually map the output of proximity sensors distributed on the car's bumpers to the repetition period of a single tone, going faster as the car draws near an obstacle. The successive tones eventually merge into a continuous tone when the car-to-obstacle distance goes bellow a given threshold, as a warning against the imminent collision. Apple's laptops[3] produce a short tone when one presses the +/- volume control buttons, whose loudness illustrates newly defined volume level.
A given PMSon efficiency mainly relies on three design choices: mapping dimension, polarity, and scale [119]. The mapping dimension defines the sound parameter to which data values are mapped. The polarity indicates if data and sound parameter variations are proportional (positive polarity) or inversely proportional (negative polarity). The scaling defines the transformation between data and sound parameter spaces, e.g. temperatures in [0°C, 40°C] mapped to a pitch in [100 Hz, 200 Hz] or [100 Hz, 400 Hz]. Listener's interpretation of the data will depend on these design choices [120], and so will the length of the learning process required to fully take advantage of the sonification.
The conception of a PMSon starts with a careful examination of the quantity, dimension, and range of the data to be sonified, along with a clear definition of the nature and requirement of the monitoring task (e.g. background surveillance, focused analysis, etc.). As for any sonification technique, the sound parameters employed, their range, and dynamic variation have to use the capabilities of the of the auditory system, while respecting its psychophysical limits [121].

The concept of MBSon has been introduced by Hermann et al. in [122] and further developed in [123] as an alternative to PMSon for high-dimensional data exploration. Where audification is inherently limited to specific scenarios, PMSon is a rich and versatile technique which however presents some limitations regarding the scalability and inconsistency of its application [122]. MBSon implies to directly explore data through interactions with a "sound-capable system", whose parameters (size, density, structure, etc.) are defined by said data [124]. Much like knocking on a wall to locate hollow cavities or monitoring the level of water poured in a container by listening to pitch variations, MBSon design is closely related to the way humans monitor and interact with their environment through sound.
The strength of MBSon is the combination of numerous potential modes of interaction (hit, rub, scratch, etc.) [125] with the capacities of the auditory system to detect subtle

[3]Other laptops obviously implement similar PMSons, Apple's are just more representative.

variations in complex sound patterns [126, 127]. As MBSon's full potential lies within high-dimensional datasets sonification, typical applications will usually involve a certain degree of training along with sharp listening skills.

Most sonification designs will employ several sonification techniques, alternatively applied according to the nature and complexity of the information to convey. Merging techniques to present several layers of information through a unique auditory object can also be applied to lighten an auditory scene without reducing its meaning [128].

7.1.3 Evaluation of Auditory Displays

The definition of an auditory display adopted here is similar to the one proposed in [129], i.e. a system that uses sound to present users with information related to a set of data, and potentially provides interaction paradigms that allow to refine their understanding of these information, as illustrated in Figure 7.3.

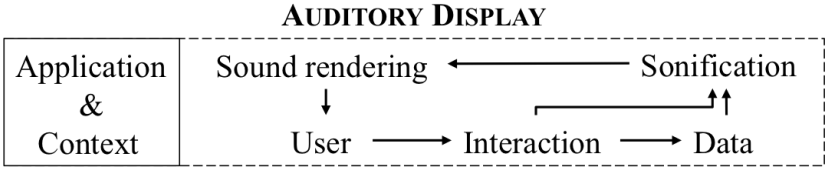


Figure 7.3: Adopted definition of an auditory display, after [129].

As for any HCI, the performance and aesthetics evaluation of an auditory display involves perceptual tests. As for any HCI evaluation, the main issue when designing these tests is to strike a proper balance between controlled laboratory experiments and real-life conditions [130, 131]. A common practice is to break down the HCI into smaller components (e.g. feature or task-wise) to reduce the impact of inter-factors relationships during the evaluation. Added to the control of the environment, stimuli and subject population, this partitioning allows one to produce meaningful and reproducible results that can be compared with other studies. The evaluation detailed in Chapter 9 provides a typical illustration of such a partitioning, defining a task abstraction to focus the evaluation on a specific perceptual issue.

As yet, nothing guarantees that these results will reflect the behavior of the whole auditory display. Conclusions issued from the listening of a twenty minute sonification over a fixed set of data in a quiet environment will not necessarily apply to its daily use e.g. by stockbrockers in a floor trader [132]. The design of an auditory display should be based on a constant back and forth between laboratory and nominal conditions involving both average and targeted populations. It is also advised to conduct evaluations on both trained experts and complete neophytes. The latter will provide an insight on system accessibility while experts will be able to assess the full potential

of a complex sonification design.

Finally, it is advised to complete any quantitative evaluation with some qualitative assessment. As pointed out by Edworthy in [133], performance outcomes and sound aesthetics are independent. The evaluation of an auditory display based on only quantitative tools could lead to efficient yet distressing or debilitating designs. This is a frequent issue in the design of navigation aids for the visually impaired, where the acceptability of a new system is commonly limited by aesthetics [134]. The more the HCI implies regular, recurrent, or prolonged use, the more the designer has to pay attention to sound and interaction aesthetics.

7.2 Literature Review: Interactive Sonification of Real-Time Data Streams

Designers obviously did not wait for the explicit definition of sonification to propose systems and interfaces based on sonic feedback. Engineers actually invented some of the most famous sound-based applications such as the Geiger counter or the SONAR (SOund NAvigation and Ranging) more than half a century before the term "sonification" appeared in the scientific literature[4].

International research efforts and coordination on auditory display and sonification design started with the creation of the International Community for Auditory Display (ICAD) in 1992. Initially, sonification was perceived "*like a solution in search of a problem*" [135] as it did not seem to confer any specific advantage over visualization in HCI design. It gradually progressed to the status of a research field in itself thanks to successful applications in interfaces for the visually impaired [136] or for parallel task monitoring where visual attention was already focused elsewhere [137, 138].

The notion of interaction, in its most simplest form being an action inducing a reaction, can be considered as a fundamental element of the most early sonification applications, as the whole function of a loud repeating sonic alarm is to incite the user to stop it. The term "interactive sonification" refers however to a user interaction that aims at refining the presented information, much like steering a Geiger counter to locate the source of a radioactive emission. During the first Interactive Sonification (ISon) workshop in 2004, Hermann and Hunt [139] defined interactive sonification as "*the discipline of data exploration by interactively manipulating the data's transformation into sound*". The assumed relationship between interactive sonification and data exploration is emphasized reading through the ISon proceedings[5] of this last decade. Data sonification provides endless case studies for interactive sonification [140], where

[4]The term sonification was previously used in chemistry to define the application of sound energy to agitate, deagglomerate, dissolve, etc. particles in a solution, now replaced by "sonication".

[5]http://interactive-sonification.org

the mere exploration of a 1D dataset requires to provide users with an interface to control their evolution in the data stream or topology [141].

As interactive sonification progressively proved its worth through simple data exploration applications [142–144], an ongoing necessity for interfaces to display larger and more complex datasets spurred some investigations on more efficient interaction and sonification paradigms. The three major research trends that emerged to meet these requirements were (1) to propose sonification strategies that increased the amount of information one could simultaneously process through sound, (2) to develop intuitive means of interacting with the data to allow the users to refine their exploration and, (3) to implement techniques reducing data dimensions with a minimum loss of significant information.

(1) Several studies have previously focused on the throughput of information that could be conveyed in a sound environment, partly related to the work of Bregman on *auditory scene analysis* in the 1980s, described as "*the process by which the auditory system separates the individual sounds in natural-world situations*" [145]. Most of these studies were based on PMSon designs where multiple variables were presented through independent auditory streams. The attribution of different pitches, timbres, or spatial positions to each auditory stream proved relatively efficient to produce perceptively "orthogonal" displays [146, 147]. One of the famous research topics related to stream segregation was the so called *cocktail party effect*, where focusing on a specific voice [148, 149] or audio stream [150] from a given subset got simpler as the sounds came from separated directions or presented significant timbral or tonal differences (e.g. the ability to focus on a specific instrument in an orchestra). Some studies on the other hand successfully increased this information throughput by mapping multiple variables onto different parameters of a single auditory stream [151, 152]. In [153], Schuett et al. for example reported the efficiency of different sound parameter mappings on listener's capacity to segregate parallel trends in weather reports.

While it is perfectly possible to map ten or more data streams to a single auditory scene by carefully designing perceptual orthogonality between sounds and amongst sound parameters, listeners will generally have a hard time to monitor more than one or two data streams simultaneously [89]. Compared to musical composition in general, the sound patterns issued from most sonification mappings are seldom designed to be perceived as a whole. As they will not correspond to any familiar composition scheme, listeners' attention can hardly focus on anything higher than individual sound events, nor will they be able easily recall melodic lines via short-term memory afterwards [154].

(2) Regarding the development of interactions in auditory displays, the exploration of data through *auditory graphs* appeared at first as a perfect application to design and test new paradigms. Basic interaction features such as pan, zoom, viewpoints selection,

etc. were already well defined for visual displays. Their adaptation to auditory graphs represented a primary basis to ascertain the usability of interactive sonification for concrete data exploration applications [155,156]. Researchers were however confronted and have been focused since on issues related to the evolution through the data stream more than on the interaction itself, and the call from Flowers [157] and Stockman et al. [158] for "*auditory viewpoints*" control through natural interfaces in auditory graphs remains yet to be answered.

Most design issues in auditory graph exploration were related to the fact that the ear does not possess the eye's ability to ignore everything but the object or region it is focused on. "Focused hearing" was achieved by sonifying only a part of the graph, windowing the data to avoid overcrowded and confusing auditory scenes. Interaction paradigms developed for auditory graphs were essentially dedicated to the positioning of this window or to the control of its automatic panning speed through the dataset. Efficient designs generally involved tactile [159] or haptic [160] interfaces were users would define their auditory viewpoint in the graph by selecting a coordinate in the cartesian space of the interface, usually mirroring all or part of the data distribution.

This mirroring of the data distribution/topology by the control interface was progressively democratized in interactive sonification design. Most common applications concerned maps or navigation aids designed for visually impaired individuals, where some sensorial substitution apparatus would reproduce the data topology, most of the time involving spatial hearing [161,162]. As for auditory graphs, the objects taken into account by the sonic restitution were usually narrowed to a user-defined region, e.g. by aiming the apparatus towards a specific region of the exploration space [162–165].

In parallel of these investigations on auditory graph interactions, researchers initiated reflections on the design of sonic feedback to inform users on their interaction context. Several studies assessed the benefits of sound-based labels, ticks, or references; added to the main data sonification to provide contextual information during the exploration [166, 167]. Generally based on earcons, auditory icons, or spearcons, these contextual information proved essential to assist and comfort users through their interactions [168], from the simplest "button pressed" acknowledgment to more complex "current position" or "current zoom/gain level" sonifications.

The notion of interactive sonification took on a whole new dimension with the introduction of Model Based Sonification in 1999 (see Section 7.1.2). MBSon is based on the simple observation that humans are used to gathering information on everyday objects by simply listening to the sound they make when subjected to specific interactions [169] (e.g. tap on a fruit to assess its ripeness). The richness of this technique came from the multiple combination of interaction-producing-audio-event feedback one could design [122,125], the interface presenting different sound-based information when squeezed, scraped, hit, etc. Barrass latter introduced the concept of Acoustic Sonification (ASon) [170], an approach similar to MBSon yet based on the creation of real

objects whose shape, structure, etc. are defined by the set of data to explore. He latter illustrated the ASon concept at ICAD 2014 with the "hypertension singing bowl" [171], a 3D-printed steel bowl which structure reflected a year of blood pressure recordings, that produced sounds when rubbed or hit with a timbre specific to the subject's health status.

One of the telling illustration of the immense potential of both MBSon and ASon techniques, while not exactly MBSon nor ASon in itself, is probably the ability with which visually impaired individuals use echolocation to gather information on their environment [127]. By carefully shaping their interaction (*click* sound produced with the mouth) in terms of loudness, frequency, direction, directivity, etc. they manage to selectively probe nearby objects, surrounding buildings, or distant landscapes with startling efficiency [172, 173].

(3) Reducing the data dimensionality prior to their display has long been a common practice in visual HCI, i.e. not specific to sonification design. The main issue of such factorization has always been to strike a proper balance between presented high level data simplicity and original raw data integrity [124]. Researchers addressing the sonification of EEG signals often faced this issue, usually dealing with complex and massive amounts of signals. Based on end user requirements, their first choice in the sonification design is to select the nature of the displayed information, from raw measured EEG signals to high level preprocessed parameters, e.g. related to specific brain states [174–176].

The particularity of data reduction for auditory display is that the considered application or concept has generally already been developed as a visual display. Hence, conventional reduction tools and process are somehow associated with these applications, already acknowledged by years of applications in the visual domain. Sonification designers have to question whether or not these conventions suit auditory display, i.e. whether or not audio would benefit from rawer or higher level data presentation. For example, the analysis of statistical values admittedly handled through boxplots or correlation values comparison in the visual domain can be sonified as such [177] or redesigned for audio-based analysis [178].

Regarding the application considered in this manuscript, the real-time constraint will eventually limit any intense or non-causal pre-processing scheme and factorization, often limited in such context [132,179,180]. As the foreseen audio direction finder shares many of the attributes of navigation aids, the data topology mirroring evoked earlier will be of use, as will the sonification of contextual information and the techniques to design orthogonal sonification mappings. Real-time data acquisition will potentially raise issues related to unknown data ranges, limiting the straightforward application of the recommendations in [158, 181] on "*Identification of absolute values*" by providing listeners with references and consistent sonification mappings (i.e. always the same value

of a sound parameter for a given value of the data). As will be discussed in Chapter 9, the application context also involves data instabilities related to signal measurement noise. Not often addressed in the sonification literature, data noise or jerk reduction is usually achieved through straightforward filtering [182] when not actually used to improve user's awareness of a given system's sate [183].

7.3 A Word on the Design of Interactive Sonification

This section succinctly exposes some considerations regarding the design of an interactive sonification aimed at data exploration.

- ➙ Define end-users/application requirements. If they do exist, identify the performance, flaws, pros and cons of peer systems. As for any HCI, the design of an auditory display requires a plain understanding of the targeted application context, the information to convey, and the interactions required.
- ➙ Question the level of information (resp. data factorization) conventionally displayed (resp. applied) in peers systems. Is it optimum (too much, not enough) regarding audio-based exploration?
- ➙ Carefully select the software(s) used for sonification prototyping. The selected framework should allow the use of different sound design techniques (concatenation, parametric synthesis, etc.). Sound design is time consuming, do not hesitate to discard sound aesthetics prior to the validation of design functionality.
- ➙ Select and merge sonification techniques according to the nature and complexity of the information to convey. Do not hesitate to use redundancy to highlight essential information (e.g. two sound parameters to encode a single information), employing"*integral*" subsets of sound parameters [184]. Make sure that the sound hierarchy in the final auditory scene (which stand out, which are correlated, etc.) reflects the degree of importance and the nature of each information.
- ➙ Carefully balance the density of the final auditory scene. Rely on orthogonal sound perception, factorize related information into single sound events, etc.
- ➙ Provide sound-based feedback on users interactions and exploration context. Users should be aware of the impact of their interaction on the system's state (e.g. punctual earcons, change the overall timbre of the auditory scene, etc.).
- ➙ Attempt to offer some flexibility in the sonification design (see e.g. [134]), the interactions, the auditory viewpoints, etc. While flexibility rapidly increases the duration of any design evaluation, it may also increase users' engagement and acceptation of the HCI.
- ➙ Systematically complete lab-based (i.e. controlled) assessments with extensive tests in real~~istic~~ conditions.

7.4 Conclusion

This chapter exposed the fundamentals of sonification, with a focus on the notions related to the auditory display considered in this manuscript. The main terms related to sonification design were introduced, along with the most commonly employed sonification techniques followed by a discussion on the nature of information they were best suited to convey. A literature review on interactive sonification was proposed, concerned with research related to real-time exploration of data streams. Finally, Section 7.3 suggested some good practices related to the design of a interactive sonification.

Chapter 8

Potential DF Outputs Sonification: Design and Evaluation

Contents

This chapter details various sonification designs implemented for several levels of Direction Finder (DF) outputs. Prior to any DF prototype, the objective is to evaluate the capacity of a sound-based interface to deal with (i.e. relay the information related to) the different types of data presented in Figure 8.1, from raw sampled signal to estimated Direction Of Arrival (DOA). The sonifications presented in Section 8.3 for example have been designed based on the assumption that the DF user could be more efficient during the search if provided information on the position, level and width of three of the main peaks in the spatial power spectrum.

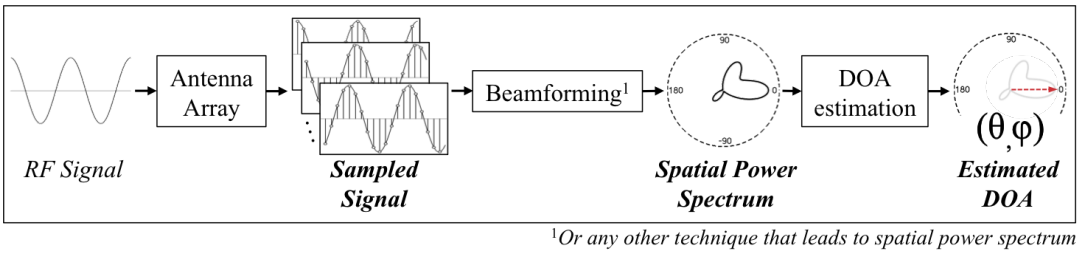


Figure 8.1: Potential outputs of a DF system. The information to sonify changes with the considered output: Sampled Signal, Spatial Power Spectrum (1D, 2D topology), or Estimated DOA (1D, 2D direction in spherical coordinates).

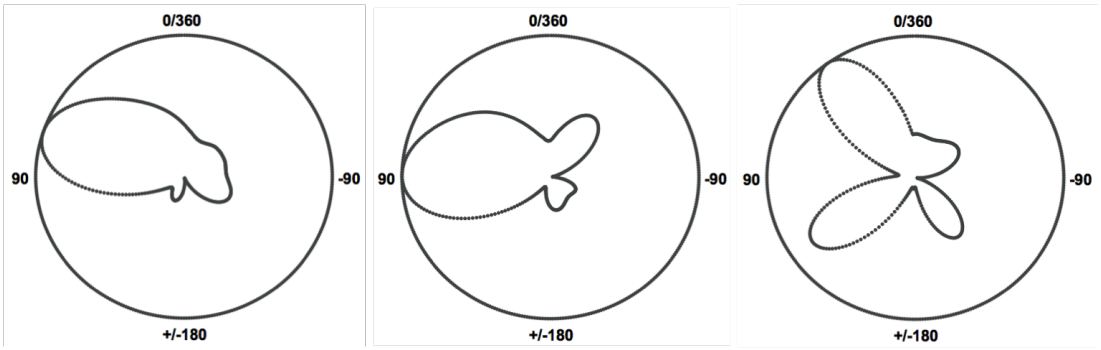


Figure 8.2: Three 2D spatial power spectrum stimuli used during the sonifications assessment. (a) Single maximum scenario, represents simple conditions in e.g. a flat field environment. (b) Two different local maxima, represents medium conditions in e.g. an urban outdoor environment. (c) Three different local maxima, represents hard conditions in e.g. an indoor environment.

8.1 Method and Experimental Design

A set of three spatial power spectrum were simulated via IlmProp, the raytracing propagation model introduced in Chapter 4. Each spatial power spectrum represented a different complexity level in the estimation, from simple (single maximum) to hard (three local maxima), illustrated in Figure 8.2. To avoid subjects detection of any experimental pattern, 9 version of each of the three spatial power spectrums were created through a simple circular shift, linearly distributed from 0° to 360°.

Two subjects participated in the experiment (1 woman, 1 man), both aged 25. For each sonification metaphor, subjects were presented with random sequences of approximately 40 spatial power spectrums each. They were asked to point in the different directions they perceived as maximums', and if possible to respect the perceived hierarchy (pointing first at the main maximum, then the second, etc.). As the objective was to assess both the precision and the efficiency of each metaphor, subjects had only 4 seconds to listen/point per spatial power spectrum before an automatic shift of the sequence to the next stimulus.

The sonification metaphors were implemented in MaxMSP[1] or PureData[2], two audio design softwares. Tests were conducted twice per subject, once with a headset and once with a bone conduction headset (also known as *bonephone*), to evaluate its impact on the subject's performance as it was seen as a potential solution for an "ears free" interface. If not mentioned otherwise, their is no impact of the headset type on

[1] https://cycling74.com
[2] http://puredata.info

perceived information in the following discussions.

The results discussed below are not issued from formal measurements, but rather from subjective assessment of participants performance by an external examinator. These listening tests were intended as a pre-evaluation of the different sonification metaphors performance. As a future study, this experimental design could be extended (more subjects, quantitative logs, etc.) to allow rigorous comparison of the considered sonification metaphor.

8.2 Sonification of an Estimated DOA

The DF outputs a unique estimated DOA $\{\theta_0, \varphi_0\}$, i.e. an estimation of the target azimuth θ_0 and elevation φ_0 with respect to the DF position (no distinction between stimuli condition in Figure 8.2). For the first sonification metaphor "DOA to binaural coupled with head tracking", the spatial power spectrum stimuli were replaced by random 2D DOA values. This sonification was the only one assessed with 2D DOAs.

DOA to binaural coupled with head tracking

Data In:	Single 2D DOA $\{\theta_0, \varphi_0\}$.
Based on:	PMSon (sound position).
Figure:	*None*
Video:	*None*
Description:	Binaural spatialization of a single sound source at location $\{\theta_0, \varphi_0, r = 1 \text{ m}\}$ coupled with head tracking.

The sonification of spherical coordinates $\{\theta, \varphi, r\}$ perfectly fits a Parameter Mapping Sonification (PMSon) based on spatial hearing, discussed in in Chapter 7. Ideally, such mapping gives the impression that the sound heard in the headset comes from the direction $\{\theta_0, \varphi_0\}$. The tests were conducted on three different binaural engines SPAT[3], LSE (binaural engine developed at LIMSI) and CW_binaural [185]. Results reflected the previous researches on the subject (see Chapter 7 and [186]), meaning that location performances were optimum for:

- short bursts of white noise, sounds with a wide spectrum and sharp attacks,
- HRTF perceptively similar to listeners',
- tracked head movements, i.e. update of $\{\theta_0, \varphi_0\}$ as listeners' head moved.

In these conditions and given enough time, we observed that trained listeners were able to estimate the correct source DOA at each iteration. Note that the few ambiguities

[3] http://forumnet.ircam.fr/fr/product/spat/

related to usual front-back or up-down confusions with non-individual HRTF [106] completely disappeared with tracking and training on a given set of HRTF, similar to the results in [101, 187]. Their exists several perceptual HRTF selection methods [188, 189] along with Interaural Time Difference (ITD) adaptation based on anthropometric measurements [190] (e.g. head radius) that can be used to reduce the length of the learning sessions. Careful design of HRTF selection, adaptation, and training with a given sound source allows one to avoid individual HRTF measurements, considered as the optimum solution for binaural rendering yet not applicable in general applications as the process requires lengthy recordings in an anechoic chamber.

The use of bonephones led to similar performance in terms of DOA estimation accuracy. As shown in [191], their potential impact on localization performance does not represent an issue for navigation based on spatial hearing. Depending on the bonephone quality and frequency response, the sonification may have to be adapted as to sound alike through both headset and bonephone interfaces. Their is no imperative need for Bone Related Transfer Function (BRTF) [191] measurements for the considered application, while a re-selection of HRTF based on perceptive tests may be advised between headset and bonephone.

Without head tracking, front-back and up-down confusions clearly hindered listeners in their estimation of discrete DOAs based on binaural sound. Following are sonification designs proposed to address this issue for a DOA constrained in the horizontal plane.

DOA to binaural SONAR scan

Data In:	Single DOA θ_0.
Based on:	PMSon (sound position), earcon.
Figure:	Figure 8.3
Video:	//assets/partII/1.1 sonification-2DpanningRadar.mov
Description:	Binaural spatialization of a single sound source at location $\{\theta_0, r = 1 \text{ m}\}$ coupled with a time based front-back discrimination feedback.

To avoid front-back confusions, this sonification added a time reference in-between DOA sonifications of θ_0 to differentiate front-back positions. As illustrated in Figure 8.3, the time reference consisted of two spatialized pink noises ($pink_L$ and $pink_R$) that periodically spanned azimuth values $\theta_L \in [0°, 180°]$ and $\theta_R \in [0°, -180°]$. When one of θ_i values equalled the current DOA θ_0, an earcon (*ping* sound) was played, spatialized at θ_0.

The time at which the *ping* was played compared to the mental representation of both $pink_i$ panning from front to back helped to avoid front-back confusions. While this sonification did remove front-back confusions, it added a delay between estimations

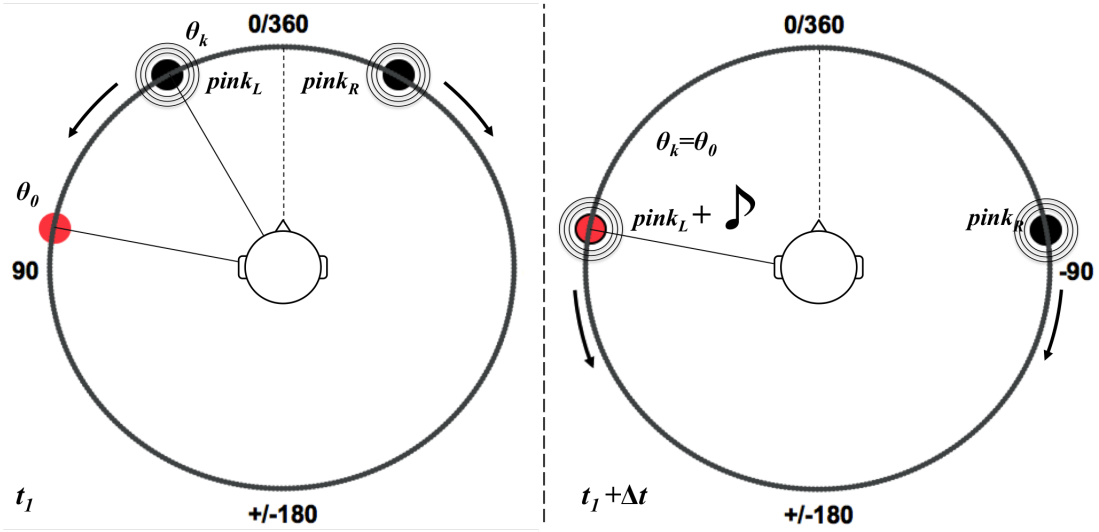


Figure 8.3: Illustration of the *DOA to binaural SONAR scan* sonification. Two sources ($pink_L$ and $pink_R$) periodically span left and right auditory space, for $\theta_L \in [0°, 180°]$ and $\theta_R \in [0°, -180°]$. When one of these sources reach the target's DOA θ_0, a short sound is produced at θ_0. The periodical span adds a time reference that prevents front-back confusions in the absence of head tracking (e.g. if the short sound occurs during the first half of the span, $\theta_0 \in [-90°, 90°]$).

and was argued to be tiresome after a few minutes of listening.

DOA to binaural arrow

Data In:	Single DOA θ_0.
Based on:	PMSon (spatial location), earcon.
Figure:	Figure 8.4
Video:	//assets/partII/1.2 sonification-2Darrow.mov
Description:	Single sound source spatialized (binaural) at location $\{\theta_0$,r $\in [-r_0, r_0]\}$ coupled with a time based front-back discrimination feedback.

This sonification was designed to give the impression of an audio compass, employing a sound that went through the head to "point" at θ_0. The sound was a pink noise, evolving along the axis defined by $\{\theta_0, r\}$ for $r \in [-r_0, r_0]$ as illustrated in Figure 8.4. When r reached r_0, a *ping* was played at position θ_0 to "point" the arrow towards the current DOA θ_0 and the sound source started a new period at $\{\theta_0, -r_0\}$/ The *ping* also served as time reference to dissociate between spanning periods.

As with the *DOA to binaural SONAR scan* metaphor, front-back confusions disappeared as the subjects learned to identify how different a given sound sounded when

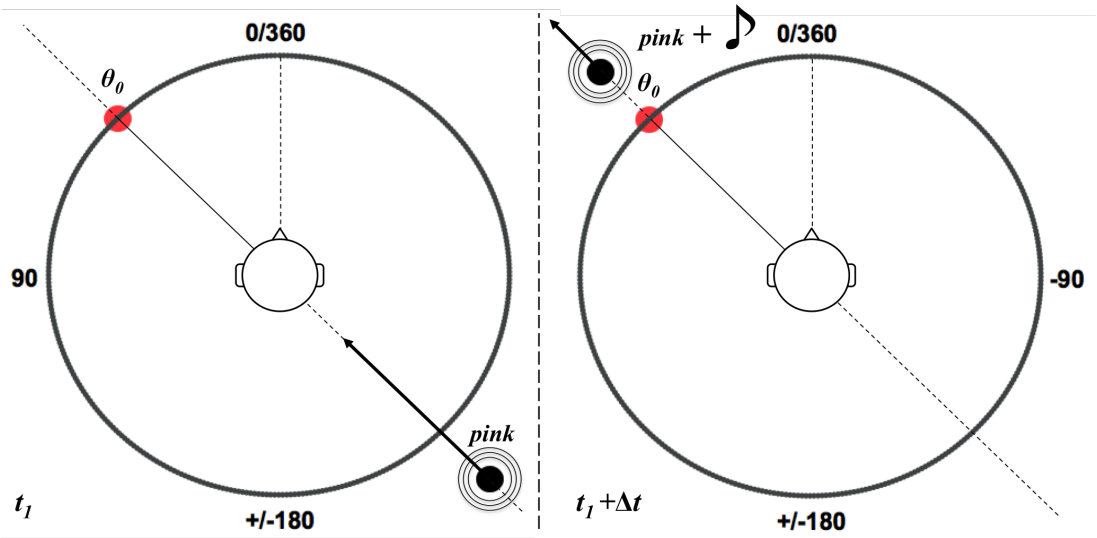


Figure 8.4: Illustration of the *DOA to binaural arrow* sonification. One source (*pink*) periodically spans the audio space for $r \in [-r_0, r_0]$ with a constant azimuth $\theta = \theta_0$. When $r = r_0$ (i.e. at the end of each span period), a short earcon is played at position θ_0. Repeated listening to front-back and back-front trajectories progressively enables to discern differences in timbre evolution between both, which quickly evolves in a mental mapping associating a certain timbre evolution to an hemisphere (front or back) for a given sound and set of HRTF.

moved from back to front or from front to back. While this sonification completely removed front-back confusions after a few minutes of training on a fixed set of HRTF, it was criticized because of its tiresome contextual sounds and the delay it imposed in between successive estimations.

DOA to binaural coupled with front-back filter

Data In:	Single DOA θ_0.
Based on:	PMSon (spatial location), earcon.
Figure:	*None*

Video:	//assets/partII/1.3 sonification-2Dspatialization.mov (without filter)
	//assets/partII/1.4 sonification-2DspatializationFiltered.mov
	(with filter)

Description:	Single sound source spatialized (binaural) at location $\{\theta_0, r = 1$ m$\}$ coupled with a timbre based front-back discrimination feedback.

This sonification simply added a low pass filter on the spatialized sound for DOA values behind the listener. Subjects clearly preferred this paradigm over both *SONAR*

scan and *binaural arrow* as it had the advantage of completely removing front-back confusions without adding any delay nor contextual sound.

These three metaphors employing sonification techniques to compensate for poor front-back discrimination in static binaural rendering could eventually be adapted for 3D spatialization, e.g. adding contextual cues or filtering specific to each DOA quadrant (to avoid up-down confusions as well). If available, a solution based on real-time head tracking would avoid the need to add contextual cues involving learning sessions and/or overloading auditory scenes.

8.3 Sonification of Discrete Values of a Spatial Power Spectrum

The DF outputs the positions of the main peaks (maximums) of the 2D spatial power spectrum, i.e. the main estimated DOA plus secondary peaks locations, otherwise dismissed by the estimation algorithm. Listening to the following sonifications, subject's first objective would be to estimate maximums positions in the spatial power spectrum topology. The relative amplitudes of the maximums and their associated widths were also considered for sonification processing. While the long term objective was to provide users with as much information as possible on the received spatial power spectrum, the following sonification rather assessed the amount of information that could be conveyed through sonification for the task at hand.

Discrete spectrum to binaural Geiger coupled with front-back filter

Data In:	One to three main maximum locations and amplitudes.
Based on:	PMSon (spatial location, amplitude, and timbre).
Figure:	*None*
Video:	//assets/partII/2.1 sonification-2Dgeiger.mov
Description:	One to three sound sources spatialized at locations $\{\theta_i, r = 1 \text{ m}\}$ coupled with a timbre based front-back discrimination feedback. Geiger counter sonification metaphor of maximums' amplitude.

As for the *DOA to binaural coupled with front-back filter*, this sonification mapped maximums location to sound sources (pink noise) location and avoide front-back confusions with a bimodal frequency filter (low pass filter for DOAs in $[-180°, -90°]$ and $[90°, 180°]$). The amplitudes of the maximums were mapped to the repetition period of a Geiger counter metaphor applied to each sound source (amplitude modulation based on a square wave). Three different sets of band-pass filters were applied on each pink noise to differentiate them, a given set always corresponding to a given maximum position in the hierarchy (main, second and third).

With a few seconds of training, subjects were able to systematically identify the main lobe and estimate its position. The fastest Geiger counter appeared to draw subject's attention with a minimum cognitive effort. At the beginning, subjects were actually able to completely ignore sounds related to second and third maximums to focus on the first maximum. After a few minutes of training, they were both able to differentiate, sort (i.e. establishing a hierarchy) and position each maximum. They judged however that prolonged listening should not involve more than 2 maximums, as it can prove tiresome to continuously monitor three or more different sound sources, particularly since in practical conditions the spatial power spectrum evolves constantly: maximums amplitudes and position can shift more than 10 times per seconds.

Discrete spectrum to binaural Geiger plus front-back filter and flanger

Data In:	One to three main maximum locations, amplitudes and widths.
Based on:	PMSon (spatial location, amplitude, timbre, pitch).
Figure:	*None*
Video:	//assets/partII/2.2 sonification-2DgeigerFlanger.mov
Description:	One to three sound sources spatialized at locations $\{\theta_i, r = 1 \text{ m}\}$ coupled with a timbre based front-back discrimination feedback. Geiger counter sonification metaphor of maximums' amplitude. Period of a flanger controlled by maximums' width.

Much like the precedent, this sonification furthermore superposed a flanger effect to each pink noise to inform subjects on the related lobes width. As the lobe grew larger, the flanger effect was accentuated to indicate the imprecision on the indicated DOA (the oscillating period of the flanger was inversely proportional to lobes widths). This sonification presented a maximum of 9 different parameters for the scenario of Figure 8.2c ($3 \times 3 \times 3$ positions, amplitudes and width).

After a few minutes of training, listeners were able to classify lobes' width as "large" or "narrow" yet none went further than this 2-states assessment. If ignored, the flanger effect was completely transparent, performance regarding maximums positioning and identification matched those of the *binaural Geiger coupled front-back filter* sonification.

Discrete spectrum to binaural coupled with echo radar

Data In:	One to three main maximum locations and amplitudes.
Based on:	PMSon (spatial location, amplitude, timbre, pitch).
Figure:	Figure 8.5
Video:	//assets/partII/2.3 sonification-2DechoBlossomDiscrete.mov
Description:	One to three sound sources spatialized at locations $\{\theta_i, r = 1 \text{ m}\}$. Maximums' amplitude mapped to time of occurrence of the sound in a periodic sequence.

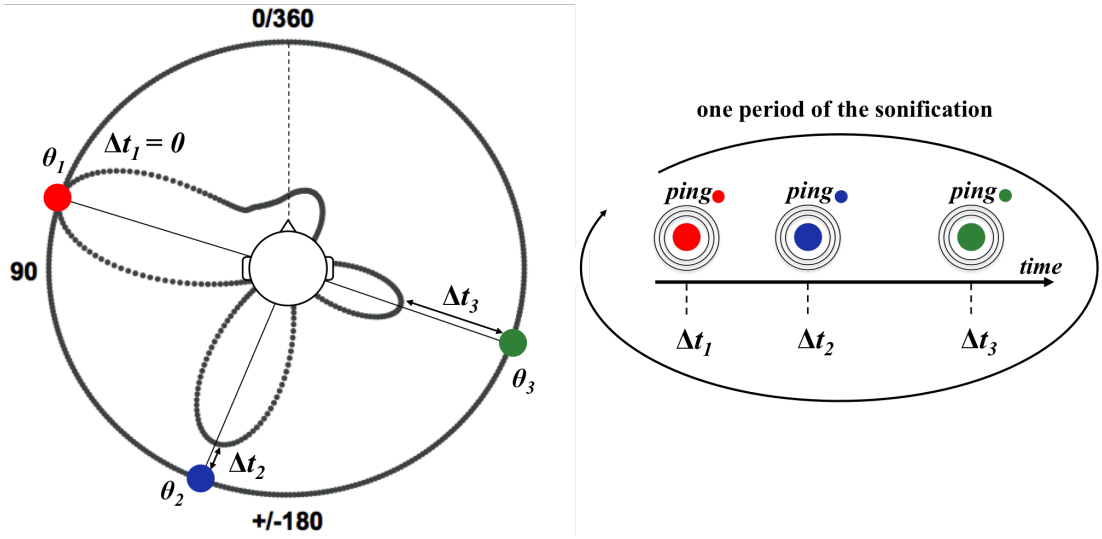


Figure 8.5: Illustration of the *Discrete spectrum to Binaural coupled with echo radar* sonification. The one to three maximums are represented by spatialized sounds, all played once over a time period. The position of a sound in time during the period is inversely proportional to its related maximum value (the higher the earlier). The naming echo radar is related to the image of a circle centered on listener's head which radius would go from 1 to 0 every period, triggering the different maximums as it crosses them.

In this design, each maximum was represented by a spatialized earcon, different for first and secondary maximums. Sound sequences were looped, beginning with the sound related to the main maximum. Second and third sounds were triggered more or less earlier in the loop based on their amplitudes relative to the first maximum: the higher the maximum, the earlier its corresponding sound in the sequence as illustrated in Figure 8.5.

This sonification did not implement any feedback on front-back positioning of the sounds sources, which resulted in uncertainties on the estimations nearly half of the time during the experiment. As for the sonifications introduced in Section 8.2, subjects judged that the time-based sonification (of maximums' amplitude) imposed a delay between successive estimations. When not certain of their estimation, subjects had to wait a complete sequence to check their first impression. While maximums' hierarchy was easily perceived, subjects had difficulties in estimating their relative amplitudes. Rather, they classified the maximums as "important" or "non-important" based on their position in the time sequence.

Discrete spectrum to binaural coupled with improved echo radar

Data In:	One to three main maximum locations, amplitudes, and width.
Based on:	PMSon (spatial location, amplitude, timbre, pitch).
Figure:	Figure 8.5
Video:	//assets/partII/2.4 sonification-2DechoBlossomDiscreteHybrid.mov
Description:	Similar to the *Discrete spectrum to binaural coupled with echo radar* coupled with a timbre based front-back discrimination feedback. Constant feedback on the main maximum's location.

This sonification was an attempt to improve the previous design by attaching a constant oscillating pink noise to the main maximum, to always have a feedback on its position. A band-pass filter was applied on the sound sources to avoid front-back confusions, as described in Section 8.2. Discrete oscillating pink noises were added to the sounds associated to second and third maximums, as the effect of the front-back filter was not perceived on the original "ping" earcons. As for the *Discrete spectrum to binaural coupled with echo radar*, maximums' amplitude were mapped to the occurrence of the earcons in the sequence. Maximums' width was mapped to the duration of each earcon.

Except for reduced front-back confusions, subjects' performance with this sonification were equivalent to the *Discrete spectrum to binaural coupled with echo radar* sonification. Same as maximums values, only two states of maximums width were perceived, namely "large" or "narrow".

8.4 Sonification of a Complete Spatial Power Spectrum

The DF outputs a complete 2D spatial power spectrum as discussed in Chapter 3, i.e. before the actual DOA estimation. The information to convey is no longer a single direction but rather a 360° topology that indicates the received power values from different directions on the horizontal plane. A successful sonification should allow a listener to be able to identify where the main signal comes from, and where are positioned other local maximums.

Complete spectrum to binaural wheel

Data In:	Complete spatial power spectrum.
Based on:	PMSon (spatial location, timbre, pitch), earcon.
Figure:	Figure 8.6
Video:	//assets/partII/3.1 sonification-2DrotatingSpectrum.mov
Description:	Unique sound source spatialized at location $\{\theta \in [0°, 360°], r = 1 \text{ m}\}$ (revolving around the head), filtered by a band-pass filter with center frequency controlled by spatial power spectrum value in θ.

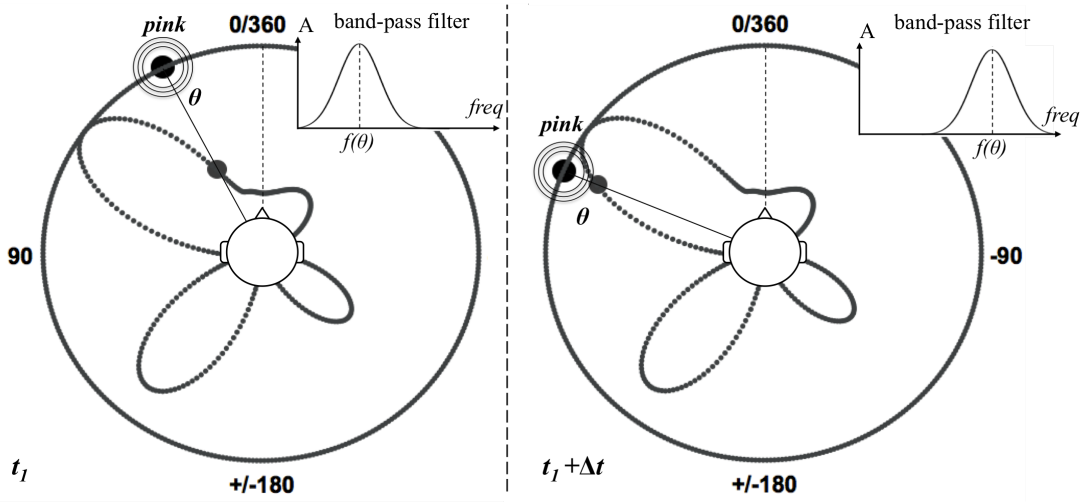


Figure 8.6: Illustration of the *Complete spectrum to binaural wheel* sonification. A single source (*pink*) periodically revolves around the user for $\theta \in [0°, 360°]$, filtered by a band-pass filter of center frequency $f(\theta)$ that increases with the spatial power spectrum value. A short earcon is played for $\theta = 0°$ to mark the beginning of a new period.

This sonification consisted of a pink noise constantly revolving around listeners' head. The perceived pitch of the pink noise (filtered by a band-pass filter) reflected the value of spatial power spectrum for $\theta \in [0°, 360°]$.

While functional, subjects complained that this sonification required to wait for a complete rotation cycle before making any decision on the correct DOA, even when the rotation speed was increased after a few minutes of training. Subjects reported that an explicit illustration of the revolution pattern (clockwise or counter-clockwise) helped to remove front-back confusions and facilitated the creation of a mental representation of the spatial power spectrum.

8.5 Conclusion

This chapter presented some of the sonification metaphors designed to convey the information related to potential outputs of a classical DF. The considered outputs ranged from raw spatial power spectrum to estimated DOA. Prior to any DF prototype construction, it remains unknown which output level leads to the best performance. Users may be more efficient with a complete 360° topology of received signal strength than with a simple unique DOA of uncertain validity.

A pre-evaluation was proposed, based on a set of two subjects asked to use the various sonifications to estimate DOAs and features in the power spatial spectrum (max-

imums' location, amplitude, etc.). Subjects were presented with different propagation scenarios and sonification metaphors. The simplest subset of scenario/sonification supposed the estimation of a single DOA, the hardest the estimation of 9 different parameters (3 maximums in the spatial power spectrum along with their relative amplitudes and widths).

The objective was to informally assess the capacity of various sonifications to convey the information related to each scenario and to identify good practices which could be suitable for the task under consideration. Future work would involve the design of a complete experimental protocol to formally assess the performance of each sonification and the number of parameters that can be simultaneously monitored by untrained/trained listeners.

The sonification of a single DOA did not represent any difficulty. For an antenna array design as presented in Chapter 5, the sonification would benefit from head tracking, removing most front-back and up-down confusions. Where the perceived sound location seemed an obvious candidate to represent the estimated DOA, other parameters (pitch, tempo, etc.) could be employed to convey information on e.g. received signal strength or measurement uncertainty. Different sound timbres can be used to easily differentiate front-back and up-down locations to further assist beginners in their listening task.

Regarding the sonification of several potential DOA candidates issued from the spatial power spectrum, most of the designs that used more than two sound sources were perceived as quite demanding, particularly for prolonged use. As yet, the evaluations presented in this chapter were not based on real-time data acquisition. It is likely that the introduction of measurement noise or fast fluctuations in the spatial power spectrum will further restrict the number of DOAs that can be simultaneously monitored. Some of the proposed paradigms which introduced additional time delay between successive estimations are to be avoided (e.g. *DOA to binaural SONAR scan*), as it is foreseen that users will need a constant low-latency feedback during their exploration, at the very least regarding the information related to the main DOA to follow.

The sonification of the complete spatial spectrum, as opposed to discrete parts of it, raised a serious time-related issue. So far, we could not identify a sonification paradigm that conveyed a complete 360° (or worse $360^\circ \times 180^\circ$ for azimuth and elevation) topology without introducing some panning period that limited the listeners in their estimation speed. Attempts at simultaneously sonifying the complete spatial power spectrum, for example mapping its discrete values every 10° to the pitch of spatialized sound sources resulted in auditory scenes judged as too complex, requiring too much concentration to barely achieve the performance obtained with the sonification of the three main maximums only. An in-between solution, considered in Chapter 9, is to provide users with a complete topology but with an interface to focus their auditory viewpoint on a

specific region.

An aspect of realistic conditions discarded by the task abstraction is the dynamic increase of the measured power range (from max to min) as the DF progresses towards the target. Prior to any navigation tests, the PMSons proposed in this chapter will need to be modified to take into account the variations in Received Signal Strength (RSS), for users to perceive when they are actually progressing towards the target. Chapter 12 discusses this RSS sonification whose main issue is related to the sensitivity of the PMSon scaling that needs to be adapted to the inverse square law dynamic followed by the RSS.

Good practices and guidelines issued from this informal study on DF output sonification metaphors and paradigms are listed hereafter. As this investigation lacked some formal perceptive tests, these remarks and observations are based more on discussions and personal observations rather than scientific results.

- ➙ While short bursts of white noise suit binaural perception, prefer a pink noise modulated in amplitude by a sinusoid waveform. This will reduce sonification harshness with a minimum impact on location accuracy.
- ➙ Additioning several pink noises, each pre-filtered with a passband filter to form an harmonic "whisper" further enhances the perceived sound quality. Filters' frequencies can be selected to form a chord, as long as every frequency band required for binaural location is represented [192].
- ➙ Avoid auditory scenes that sound complex from the start by reducing the prevalence of secondary information sonification. Beginners must be able to effortlessly focus on the main stream of information. Training will progressively lead listeners to discern more subtle patterns, allowing them to gather more information from the sonification.
- ➙ Minimize any delay induced by the sonification. Avoid cyclic sonifications where users have to wait a whole period to confirm their estimation. Prefer sonifications that provide a constant feedback on every major parameter. Limit punctual notifications to minor contextual features (e.g. volume level feedback).
- ➙ Do not use loudness as a sound parameter. Use of loudness usually results in masking effects between sound events. Their is no such thing as an "absolute pitch" for loudness: no amount of training enables listeners to build a mental map attributing a certain loudness to a given e.g. RSS level. Finally, listeners will probably need to modify the system's volume according to the level of environmental sounds.
- ➙ Avoid unnatural PMSon polarity, such as increasing a rhythm to indicate a decrease in RSS level. Take into account what is considered as the natural polarity for the targeted population [193].

➙ For systems involving some uncertainty on maximum and minimum values, make sure to keep some flexibility in parameter mapping to avoid situations where the sonification exceeds perceptual thresholds.

Prior to the suggested formal assessment of the sonifications proposed in this chapter, we strongly encourage investigations on MBSon or Audification paradigms (rather than PMSon) for DF output sonification. Enabling the creation of much richer and complex auditory scenes, MBSon and Audification paradigms usually involve long training sessions yet raise potent designs, much like the SONAR used by submariners that often requires several years of training to master. During the thesis, we initiated some tests on the design of raw Audification of the output of a superheterodyne receiver: based on two RF antennas positioned at $\lambda_{RF}/\lambda_{audio} \times interauralDistance$ to reproduce phase-related ITD perception. These tests did not produce any functional solution, yet would deserve some more investigations. Future research could be inspired by the "*bionic ears, listening to radio waves*" study presented in [194] or the design of C. Kubisch on straightforward RF audification[4], which while approached from an artistic point of view perfectly illustrates the suggested line of research.

[4] http://www.christinakubisch.de/en/works/electrical_walks

Chapter 9

Real-Time Reduction of the Perception of Measurement Noise Based on Psychoacoustics Considerations

Contents

This chapter presents a comparative study on different Parameter Mapping Sonifications (PMSons) designed to reduce the perception of instabilities during the exploration of noisy data streams. Based on the real application of beacon localization for rescue operations, an abstraction of the task was developed which simulates the

basic concepts involved, i.e. the estimation of local maxima during the exploration of a 1D data topology. Three different PMSons were designed, based on pitch, pitch averaged over time, and tempo, evaluated for three different noise levels applied to the data. Subjects explored the auditory graph using a pen tablet to define their auditory viewpoint (i.e. position) in the topology. Evaluations were based on subjects' ability to quickly appraise the underlying data topology along with aesthetic considerations regarding the impact of noise on the exploration task. Results showed that both tempo and pitch averaged PMSon reduced the perceived instabilities compared to pitch, while tempo preserved the response time of the sonification feedback.

Part of this chapter has been submitted as a paper journal for consideration for publication in the *International Journal of Human-Computer Studies (IJHCS)* for the 2015 special issue on data sonification and sound design in interactive systems[1].

9.1 Introduction

Figure 9.1 illustrates the problem that triggered this investigation, where the use of a directional antenna based Direction Finder (DF) involves some measurement noise inherent to signal strength acquisition. Said DF can be seen as a "Geiger counter for cellphone signals", composed of a directional antenna that feeds Radio Frequency (RF) power measurements to a sonification algorithm. Operation is based on the fact that an increase of the received power occurs as the directional antenna is steered towards the RF beacon. Supposing a PMSon-based interface, a rescuers main task is to explore an auditory graph where they need to listen and compare fluctuating sound parameter values. This raises the question of the suitability of the traditional pitch-based approach [157,165,195–197] as a direct sonification of these data could lead to serious fatigue issues caused by prolonged listening to continuously *jittering* tones. The objective of this study is to evaluate the capacity of two different PMSon designs to reduce the fluctuations perceived while preserving both efficiency and accuracy of users' estimations during the exploration task.

The remainder of this chapter is organized as follows. Section 9.2 discusses the choice of a tempo-based PMSon for noisy data sonification. Section 9.3 details the experimental design: task abstraction, PMSon, input data, and topology design. Section 9.4 describes the experimental method: subjects, stimuli, apparatus, and procedure. Finally, Sections 9.5 and 9.6 present the results and the discussion while Section 9.7 exposes the conclusion.

[1] D. Poirier-Quinot, G. Parseihian, and B. F.G. Katz, "Reduction of perceived instabilities in Parameter Mapping Sonification: application to the real-time exploration of a noisy stream of data".

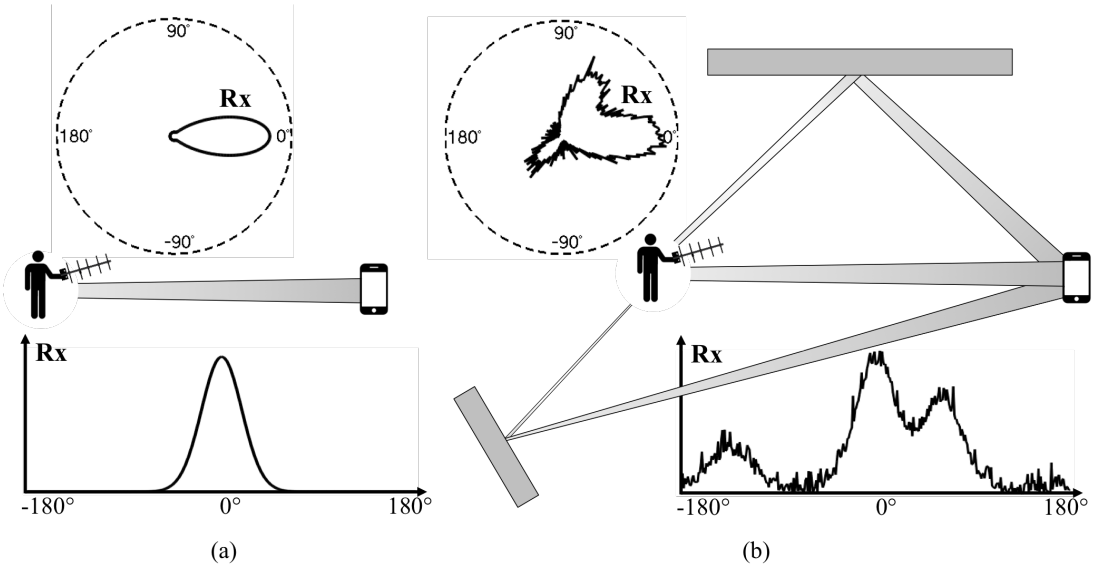


Figure 9.1: Illustration of two different propagation scenarios and their impact on direction of arrival estimation by a DF user. (a) Simple propagation scenario; the only signal that reaches the DF corresponds to the direct path towards the RF beacon. The DF user observes a single maximum in the received signal strength. (b) Propagation scenario in a more complex environment (e.g. urban area); multiple signals reach the DF due to reflections of the original signal emitted by the RF beacon on elements of the environment. Here, the DF user would have to identify left and right local maxima as reflections and focus on the global maximum of received signal strength, assuming it corresponds to the direct path towards the RF beacon. The measurement noise intrinsic to the acquisition is also represented.

9.2 Sonification of Noisy Data

Following the sonification study in Chapter 8, tests on the physical DF prototype introduced latter in this manuscript (see Chapter 11) were conducted employing a Pitch-based PMSon of the received RF power. While the design produced acceptable performance, noise induced instabilities on the pitch were reported as tiring by subjects exposed to prolonged listening. In the current study, the basic Pitch PMSon is compared to two other sonifications designed to reduce noise perception. The first implements a *running average* upstream of the Pitch PMSon, adjusted to remove most of the noise component with a minimum impact on general system latency. The second is based on *tempo*, more precisely on variations of the Inter-Onset Interval (IOI) between sound bursts as a function of DF power measurements, a sonification design often referred to as the Geiger counter metaphor.

Based on psycho-acoustic considerations, the IOI-based PMSon can be seen as a

natural equivalent of the running average. As emphasized in most perception models [97, 198], tempo perception is based on a semi-conscious evaluation of burst occurrences over time intervals. Compared to pitch, this could suggest that listener's attention will more readily ignore noise-related tempo fluctuations or irregularities and will be able to focus more on averaged estimations. Compared to computational averaging, the length of these time intervals depends on the IOI duration, decreasing as the tempo increases. This observation can be seen as a corollary of Weber's law [199] on Just Noticeable Differences (JND) of tempo variations which describes humans ability to detect temporal variations as increasing when the IOI decreases. Such an adaptive dynamic of estimation and precision could offer an advantage for the considered application.

PMSon designs will be assessed according to (1) speed and precision of subjects' estimation of the global maximum position and (2) precision of subjects' estimation of local maxima positions and relative amplitudes, for different noise levels and topology difficulty.

9.3 Experimental Design

This section provides implementation details of the task abstraction for the current experiment, data input, and sonification mappings.

9.3.1 Task Abstraction

The real rescue search task can be seen as a spherical topology exploration, where the DF user has to find the direction of arrival of the strongest RF signal. As seen in Figure 9.1, this task can be reduced to the exploration of a 1D topology if the DF user only focuses on azimuthal estimation (e.g. at distances far from the RF beacon relative to any height information). In order to ascertain condition repeatability across subjects, the abstracted 1D exploration task is based on pre-recorded DF inputs. To simplify the user interface, the exploration is carried out on a pen tablet: simulated DF antenna orientation is mapped to pen position on the tablet's horizontal axis. In sliding the pen across the tablet, subjects will hear the sound parameter variations corresponding to $0 - 360^{\circ}$ DF power measurements (see Section 9.3.2).

Note that this abstraction may slightly affect the ecological validity of the results regarding the original task which is fundamentally an unconstrained angular exploration. The abstraction of the task to a linear exploration constrained to the 20 cm tablet width, compared to the angular $2\pi\times$(arm length) arc, is likely to increase the exploration speed at the expense of the estimation accuracy and possible border effects at the tablet edges. As such, this study is constructed more as a comparative evaluation of the three PMSon designs and not an absolute measure of the expected results

in the real search condition.

9.3.2 Input Signal

The topologies used as PMSon inputs were based on measurements issued from an actual DF prototype attuned to a RF beacon emitting electromagnetic bursts at 900 MHz. Three different scenarios were considered, corresponding to different radio propagation conditions between RF beacon and DF antenna. Scenario difficulty ranged from simple to hard regarding the exploration task, composed of one to three local maxima. While scenarios 1 and 2 were based on actual measurements, the third was simulated to represent a worst case scenario in order to test the PMSons design limits in critical conditions.

Scenario C_1 **- simple**: outdoor, flat field, nearest RF reflective surface at approximately 50 m from both DF and RF beacon; DF at 30 m from RF beacon; 1 global maximum.

Scenario C_2 **- medium**: indoor, research facility, DF at approximately 30 m from RF beacon; DF and RF beacon in different rooms with 2 walls in-between; 1 global maximum, 1 additional local maximum.

Scenario C_3 **- hard**: simulated worst case scenario, crowded urban area, DF at approximately 30 m from RF beacon in different streets surrounded by buildings; 1 global maximum, 2 additional local maxima.

DF measurements of Root Mean Square (RMS) received RF power values were recorded for 8 different azimuthal orientations of the directional antenna at a fixed position. 100 RMS values were recorded every 45°, at a sample rate of 50 Hz, from 0 to 315°, with 0° orientation corresponding to the directional antenna pointing directly towards the RF beacon. Each RMS measurement was calculated from 512 instantaneous power values sampled at 100 MHz. The final basic topologies for scenarios C_1 and C_2 were obtained based on a second order spline interpolation between the 8 directional values obtained after averaging each set of 100 RMS values. The third topology was derived from a geometry-based simulation [200] in a crowded urban environment. The three resulting 0-360° topologies are shown in Figure 9.2, unwrapped on a normalized linear scale.

Three conditions were defined regarding RMS measurements noise level: **noiseless** N_0, **noisy** N_1, and **very noisy** N_2. The "noisy" N_1 condition was derived from actual RF measurements obtained in propagation conditions of scenario C_2. 100 RMS values were recorded every 45°, at the nominal audio DF sampling rate of 50 Hz. After removing any DC component, typical noise variance and spectral density were calculated and averaged across antenna orientations. Based on these parameters, a noise model was created to alter the data upstream of the sonification. The resulting model deliv-

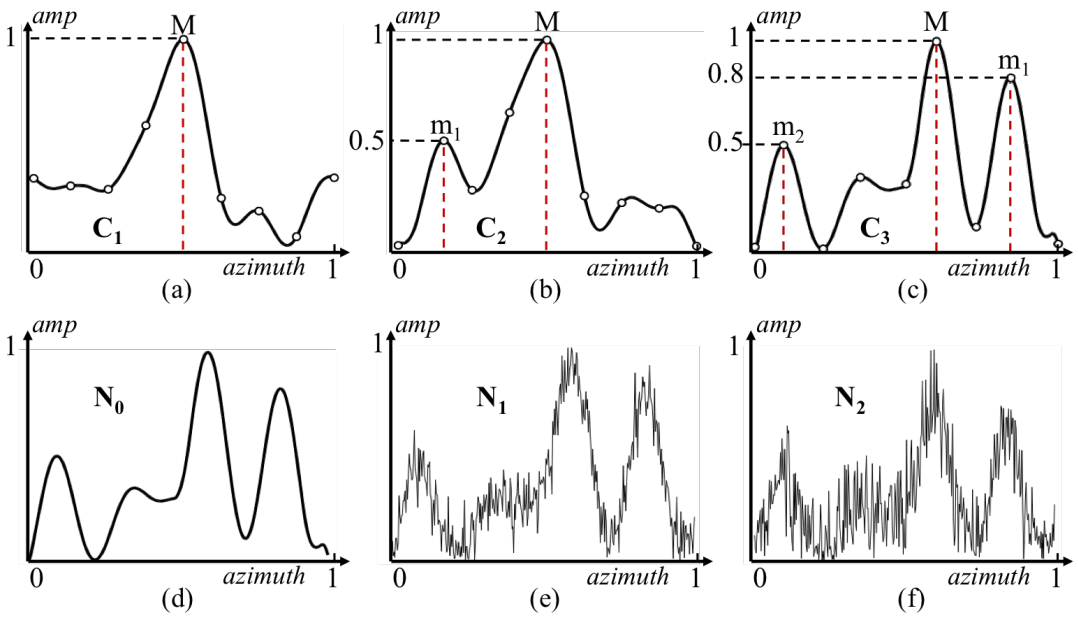


Figure 9.2: Illustration of the three data topologies and noise levels. (a-b-c) Normalized plots of the received signal power as a function of antenna orientation (azimuth) for the three scenarios simple (C_1), medium (C_2) and hard (C_3) from left to right. The ($\circ$) represent averaged values over 100 measured RMS. (d-e-f) Normalized plots illustrating the three noise level conditions noiseless (N_0), noisy (N_1) and very noisy (N_2) as applied to topology C_3. Topologies were re-normalized in amplitude after the noise addition, prior to the sonification. Noise variances were slightly augmented to result in $0.8e^{-2}$ (N_1) and $3e^{-2}$ (N_2) after normalization.

ers instantaneous values of a low-pass filtered white Gaussian noise of variance $0.8e^{-2}$, cutoff frequency 12 Hz, quality factor $Q = 0.7$, and gain=1. The update rate of these values was adjusted to 50 Hz to mirror the DF audio update rate and produce data instabilities that sounded comparable to those perceived during the recording session of scenario C_2. The "very noisy" N_2 condition was based on the N_1 condition, however with a $3e^{-2}$ variance, arbitrarily defined in order to explore the limitations of the three sonification designs.

9.3.3 Sonification Metaphor Design

A given PMSon efficiency relies principally on three design choices: mapping dimension, polarity, and scale [119]. The mapping dimension defines the sound parameter to which data values are mapped. Polarity indicates if data and sound parameter variations are proportional (i.e. *positive* polarity) or inversely proportional (i.e. *negative* polarity).

Finally, the scaling defines the bijection elaborated between data and sound parameter domains.

The objective comparison of sonification metaphors regarding their capacity to provide information on a data-set eventually leads to considerations of normalization across perceptive scales. Parseihian et al. in [195] designed PMSons for a similar study based on the Just Noticeable Difference (JND) approach of Grond & Berger [121]. This approach attempts to establish a one-to-one JND scale between sound parameter variations. As a result, listeners should have the same sensitivity when exploring data through each PMSon. We shall refer to this concept as "inter-PMSon iso-sensitivity". This concept also suggests the implementation of PMSon scaling functions based on perceptual linearity, providing "intra-PMSon iso-sensitivity". These sensible guidelines have been adopted for the PMSon designs in the current study in order to legitimize inter-PMSon comparison and facilitate inter-study cross-analysis.

9.3.3.1 Pitch (P)

Sound parameter: The pitch based PMSon P is based on the perceived frequency of a sound, here a pure tone:

$$s(t, y) = A\big(f_{pitch}(y_{norm})\big) \cos\big(2\pi f_{pitch}(y_{norm})t\big)$$

where $y_{norm} = C(x_{norm}) \in [0, 1]$ represents the normalized topology height at the pen normalized horizontal position $x_{norm} \in [0, 1]$ on the tablet, and $f_{pitch}(\cdot)$ the scaling function. The weighting coefficient $A(\cdot)$ is applied to the amplitude of the sine wave to ensure a constant perceived loudness across all frequencies, after the isophonic curve from Standard ISO 226 [96].

Polarity: positive (pitch increases as data values increase).

Scaling: The scaling function $f_{pitch}(\cdot)$ is logarithmic to match human perception (intra-PMSon iso-sensitivity):

$$f_{pitch}(y_{norm}) = f_{min} 2^{y_{norm} n_{oct}}$$

with $n_{oct} = (\ln 2)^{-1} \ln(f_{max}/f_{min})$ the number of octaves covered by the sonification and f_{min} and f_{max} the extreme frequency values (also designed as ambitus). The scaling function spans over 2.66 octaves, from 300 to 1900 Hz, both in accordance with the [62 Hz, 2637 Hz] range recommended in [181] and equal to the number of JNDs of the IOI scaling function (inter-PMSon iso-sensitivity). Based on evaluations of pitch difference limens as a function of frequency in [93], $f_{pitch}(\cdot)$ applied between these ambitus corresponds to approximately 560 JNDs.

9.3.3.2 Pitch Averaged (PA)

Exact same mapping, scaling, and polarity as with PMSon P, with the addition of a running average RMS applied over 25 samples (i.e. 0.5 sec of data stream at 50 Hz) upstream of the sonification. This 25 samples value was carefully chosen to remove most of the perceived noise with a minimum impact on sonification latency. The running average corresponds to a sinus cardinal frequency filtering with unitary gain and first zero-crossing at 1 Hz. Supposing a noiseless topology, this PMSon behaves exactly like the Pitch PMSon when the pen is moved slowly across the tablet (e.g. below 0.1 cm·s^{-1}) and smoothes more and more the topology as the exploration speed increases.

9.3.3.3 Tempo (IOI)

Sound parameter: The tempo PMSon is based on the IOI between bursts of pink noise, achieved through cycle period variations of a square wave modulation. The square wave duty cycle was fixed to 0.2 (i.e. individual burst duration = 0.2×cycle period), with the period defined by the scaling function $f_{IOI}(\cdot)$.

Polarity: negative (IOI decreases as data values increase).

Scaling: The scaling function $f_{IOI}(\cdot)$ has been designed to provide as much intra-PMSon iso-sensitivity as possible, along with an equivalent number of JND as in PMSon P. Based on measured JND of temporal variations as a function of IOI in Figure 2 of [201], $f_{IOI}(\cdot)$ is composed of two 7$^{\text{th}}$ order polynomial interpolations:

$$f_{IOI}(y_{norm}) = \sum_{n=1}^{8} a_n y_{norm}^{8-n}$$

which produce tempos ranging from 1 to 320 cycles per second (cps), for approximately 560 JNDs.

Pending a rigorous scaling methodology, inter- and intra-PMSons iso-sensitivity will be roughly assessed through result analysis. Relaxed for the task at hand, a valid inter-PMSons iso-sensitivity (initially same number of JNDs in both PMSon) would correspond to subjects being able to evaluate topology trends with equivalent accuracy with both pitch P & PA and IOI PMSons. A valid intra-PMSon iso-sensitivity (initially linear perception across PMSon) would correspond to subjects being able to correctly evaluate local maxima amplitudes relative to each other with a given PMSon.

Table 9.1: Coefficients of the polynomial function f_{IOI} that defines IOI values as a function of topology's amplitude y_{norm}. The original function was designed based on Figure 2 of [201] which reports 560 JNDs (or Difference-Limens) between 1 and 320 cps. Values of y_{norm} 0 to 1 were mapped to IOI values from 1 to 320 such that: for $n \in [\![2, 560]\!]$, $f_{IOI}(y_{norm} = n/560) = u_n = u_{n-1} + \text{JND}(u_{n-1})$ cps with $u_1 = 1$ cps and $\text{JND}(u_{n-1})$ the value of the JND in [201] Figure 2 for IOI= u_{n-1} cps. The two 7^{th} order interpolations were calculated based on the resulting points, the interval [0,1] was split to ensure well behaved polynomial interpolations.

y_{norm}	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
$[0, 0.2[$	-5.36e^7	4.81e^7	-1.87e^7	4.24e^6	-6.40e^5	7.06e^4	-5.92e^3	3.23e^2
$[0.2, 1]$	-2.07e^3	9.84e^3	-1.99e^4	2.21e^4	-1.49e^4	6.09e^3	-1.46e^3	1.67e^2

9.4 Method

9.4.1 Subjects

18 subjects participated in the experiment (3 women and 15 men), aged between 24 and 61 (mean age 31 $\pm$ 9). Six were considered as expert (musicians or actively working in the audio domain), two were left-handed. Subjects were paid volunteers (10 €), none reported any hearing loss regarding the frequency range of the experimental stimuli.

9.4.2 Stimuli and Apparatus

Subjects were equipped with a stereo open circumaural headphone (model Sennheiser HD600) and placed in front of a pen tablet with active surface dimensions 16×21 cm (model Wacom Intuos 3:6×8). The experimental session was conducted in a quiet listening room, using an interactive interface implemented using the Max programming environment[2], running on an OSX computer connected to a RME Fireface 800 soundcard. The computer screen was used to display task instructions and feedback during both experimental and training sessions. Subjects were free to adjust the volume once at the beginning of the experiment as no PMSon was based on perceived loudness.

Stimuli were synthesized in real-time according to the PMSon defined in Section 9.3.3 based on the topologies defined in Section 9.3.2. To avoid subjects learning the 3 topologies of Figure 9.2, 9 different versions of each were created through a simple circular shift, placing the main maximum at normalized positions $\{0.1, 0.2, \ldots 0.9\}$ on the horizontal axis. PMSon ambitus, i.e. lower and upper sound parameter limits, were randomly shifted to prevent subjects from learning what each PMSon sounded like at

[2] http://cycling74.com

Table 9.2: Independent and dependent variables of the experimental protocol.

Independent variables		
subject	18	random variable
PMSon	3	Pitch (P), Pitch Av. (PA), IOI
noise level	3	noiseless (N_0), noisy (N_1), very noisy (N_2)
topology (number of maxima)	3	C_1, C_2, C_3
shift (circular topology shift)	9	C_{i1}, C_{i2}, ... , C_{i9} with $i \in \{1,2,3\}$

Dependent variables
task execution times (phases 1 and 2)
main maximum estimation (phase 1)
topology drawing (phase 2)
subjective remarks and answers to PMSon-related Likert items

global maxima, i.e. for $y_{norm} = 1$. The random shift was induced upstream from the noise injection by remapping y_{norm} values from $[0, 1]$ to $[0.1 + \varepsilon, 0.9 + \varepsilon]$ with ε in $[-0.1, 0.1]$.

9.4.3 Procedure

The chosen experimental design was a repeated measures design with 3 main factors: PMSon, topology difficulty (i.e. number of maxima), and noise level. A fourth and minor factor was the topology shift, introduced to avoid subject detection of any experimental pattern (9 different shifts). These factors, along with associated evaluation metrics concerned with task execution efficiency and strategies, are presented in Table 9.2.

The experiment was divided into 3 blocks of 27 trials each. Each block concerned a unique PMSon and presented 3 repetitions of the 3×3 noise level and curve difficulty conditions ($3 \times 3 \times 3 = 27$ trials). Noise and curve conditions were semi-randomly distributed within the trials such that a given noise level or curve was never met more than twice in a row. A balanced Latin-square crossover design was applied to the 3×6 subjects for the presentation order of the PMSons to avoid carryover effects during the evaluation, the 6 potential combinations evenly distributed amongst them. Topology shifts were semi-randomly applied such that the main maximum position on the horizontal axis was never repeated more than twice in a row. Three predefined training trials were added at the beginning of each block in order to familiarize subjects

with the PMSon used in the block. After each block, subjects answered a set of 4 questions each based on a five-point Likert scale, assessing their general impressions regarding the PMSon suitability for the exploration task.

Each trial consisted of a two-phase exploration task for a given noise, topology, and shift condition. Phase 1: subjects were instructed to find the global maximum of the topology as quickly and as accurately as possible. Phase 2: subjects were instructed to thoroughly explore the same topology so as to be able to draw it afterwards. There was no time constraint for phase 2, subjects were instructed to concentrate on accuracy rather than speed. In each trial, subjects were first notified of the beginning of phase 1 with an audio-visual feedback. Subjects then had to place the pen on the tablet to start the search and validate the estimated maximum position through keyboard press or mouse button. A second notification immediately announced the beginning of phase 2, where subjects were free to explore the same line graph in the same conditions [PMSon, noise, topology, shift]. After validation (keyboard press or mouse button), they drew the perceived topology on the tablet. Drawing validation concluded phase 2 and a new trial started with phase 1, for a different noise, topology, and shift condition combination. Other than the real-time visual display of phase 2 drawing, no feedback was provided regarding subjects performance in either phase 1 or 2.

The experiment started with a training session, different from the three pre-block training trials, where subjects were introduced to the current PMSon through a similar exploration and drawing task. Subjects were free to explore a simple topology (single maximum), select one of the three noise levels and activate/deactivate the visual feedback to understand how the PMSon reacted to the topology. Each PMSon was explicitly introduced and the objectives of each task defined. Subjects naturally used their dominant hand to hold the pen during the experiment.

9.5 Results

This section presents quantitative and qualitative results analysis, based on metrics presented in Table 9.2, followed by a discussion that concerns PMSon-related performance regarding the exploration task across noise level and topology conditions. For each phase, results were averaged across the three repetitions for each subject. The significance of results has been assessed using a 3-way repeated measures ANalysis Of VAriance (ANOVA) with PMSon (3 levels), Noise (3 levels), and Topology (3 levels) as within subjects factors, with a p-value threshold of 5%. Newman-Keuls post hoc tests were employed to test specific effects and significance levels. The topology shift showed no significant impact on subjects results and was not considered in further analysis steps.

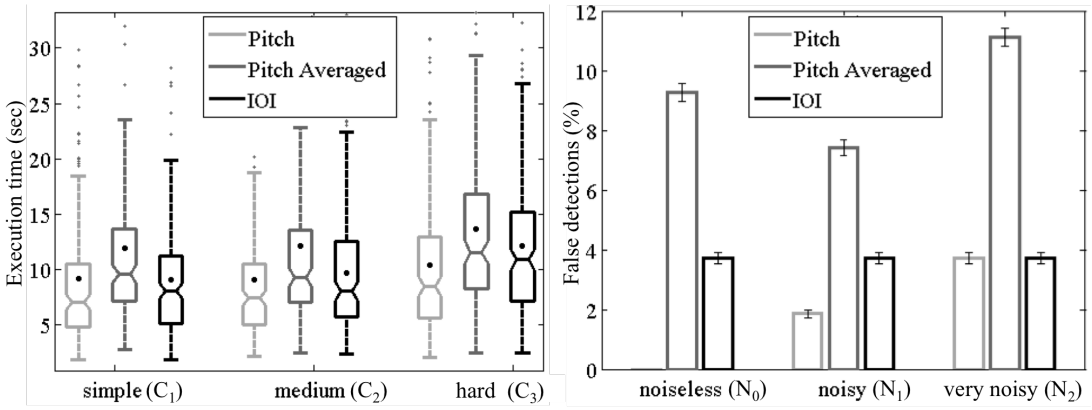
Table 9.3: Results of the 3-way repeated measures ANOVA performed the evaluation metrics with the PMSon, the Noise, and the Topology as within subjects factors (3 levels each). Evaluation metrics concern the task execution time of phase 1 and 2 (resp. Time 1 and Time 2), False detection of the main maximum in phase 1, Accuracy of the estimated maximum position, and the Correlation between subjects drawings and the original topologies. The notations F=F(2,34) and F_2=F(2,32) have been adopted for the F statistics. The reduction of 2 in degrees of freedom from F to F_2 is due to the fact that subjects expertise (2 levels) was identified as a significant factor in phase 2, but not phase 1, and was therefore disregarded in subsequent phase 1 statistical analysis. Results in sub-tables concern pairwise post-hoc comparisons (based on Newman-Keuls tests), reported only for significant ANOVAs. ε= 0.001, p-value threshold of 5%.

	Time 1		False detection		Accuracy		Time 2		Correlation	
PMSon	F=15.09, p<ε		F=3.58, p=.039		F=1.93, p=.160		F_2=8.78, p<ε		F_2=6.77, p=.003	
	PA	IOI	PA	IOI	PA	IOI	PA	IOI	PA	IOI
P	p<ε	.196	.037	.486	-	-	p<ε	.227	.545	p<ε
PA		p<ε		.069		-		.007		.002
Noise	F=17.03, p<ε		F=0.66, p=.52		F=12.45, p<ε		F_2=8.40, p<ε		F_2=1.94, p=.160	
	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2	N_1	N_2
N_0	.002	p<ε	-	-	.010	p<ε	.937	p<ε	-	-
N_1		.015		-		.030		.003		-
Topology	F=15.97, p<ε		F=15.97, p<ε		F=26.80, p<ε		F_2=0.22, p=.805		F_2=60.82, p<ε	
	C_2	C_3	C_2	C_3	C_2	C_3	C_2	C_3	C_2	C_3
C_1	.558	p<ε	1	p<ε	.144	p<ε	-	-	.066	p<ε
C_2		p<ε		p<ε		p<ε		-		p<ε

9.5.1 Phase 1: Search for the Global Maximum

In phase 1, subjects had to estimate the topologies' main maximum position as fast as possible. Reported in Figure 9.3a, the task execution time was significantly lower in both P and IOI conditions than with PA (means of 9.9 ± 6.5 vs. 12.6 ± 8.4 sec). Execution times increased with noise level and topology difficulty, significantly between each noise level condition (means of 9.1 ± 5.5, 10.9 ± 7.5, and 12.4 ± 8.3 sec for N_0, N_1, and N_2) and topology C_1 and C_2 compared to C_3 (means of 10.2 ± 7.1 vs. 12.0 ± 7.5 sec respectively).

False detection, i.e. situations where subjects designated a secondary maximum as being the global maximum, occurred only for explorations on topology C_3. This is assumed to be due to the relatively high value of the second maximum m_1 compared to M (c.f. Figure 9.2c). Figure 9.3b indicates that P and IOI caused significantly less false detections than PA (means of 1.9% and 3.7% vs. 9.3% for C_3 explorations). This is highlighted by an interaction effect of PMSon*Topology [$F(4, 68) = 5.09$, $p < 0.005$] with a significant difference between IOI C_3 and the other conditions. Noise level



(a) (b)

Figure 9.3: Results of phase 1 of the experiment. (a) Task execution time (in sec) as a function of PMSon and topology difficulty across subjects and noise levels. Black (•) indicates means. (b) Percentage of false detections as a function of PMSon and noise level across subjects for the third topology (C_3, hard). A false detection refers to subjects identifying a secondary maximum as the global maximum.

had no significant impact on false detections, despite the trend observed on P-related estimations in Figure 9.3b.

Globally, all three PMSons appeared equivalent in terms of maximum position estimation accuracy. An average estimation error of $5.8 \pm 5.2^\circ$ (equivalent to 3.4 ± 3.0 mm on the tablet) was observed, merging both true and false detections. For false detections, the error was calculated using the targeted maximum as a reference. A higher estimation accuracy was observed for topology C_3 (means of $4.7 \pm 4.1^\circ$ compared to $6.5 \pm 5.5^\circ$ for both C_1 and C_2), potentially related to the steeper curve of its main maximum compared to C_1 and C_2. The presence of noise in the data impacted the estimation accuracy, from $5.1 \pm 7.3^\circ$ (for N_0) to $6.6 \pm 9.0^\circ$ (for N_2) with a significant difference between each noise condition. The analysis of PMSon*Noise level interaction highlights a stronger dependence of Pitch PMS to noise level than the two other PMS [$F(4, 68) = 3.38$, $p < 0.05$]. Indeed, in the noiseless condition (N_0), P led to the highest estimation accuracy where as it led to the worst estimation accuracy in the very noisy condition (N_2).

The topology shift introduced in Section 9.4.2 was of no consequence regarding task execution time, false detection or estimation accuracy. The six experts were not significantly more accurate nor faster than the average, while four of them systematically identified the true maximum across every noise, topology, and PMSon conditions.

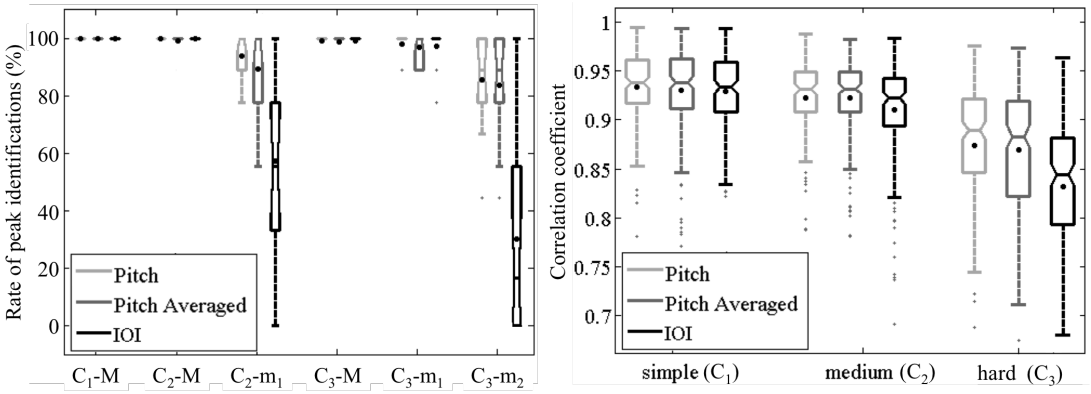


Figure 9.4: Results of phase 2 of the experiment. (a) Identification rate of each maximum in the three topologies as a function of PMSon and maximum ID across subjects and noise levels. Maxima are labelled "C_i-M/m_j", respectively the topology number (i from 1 to 3, i.e. simple to hard) and the maximum ID (e.g. C_3-m_1 is the second maximum of the third topology, c.f. Figure 9.2c). (b) Evolution of correlation values between subjects' drawings and the original topologies for the different PMSons and topologies across all subjects and noise levels. Black (•) indicates means.

9.5.2 Phase 2: Explore, Analyse, and Redraw Topologies

Phase 2 concerned subjects' ability to explore and reproduce (draw) data topologies across the different PMSons, noise levels, and topology difficulties. Results extracted from subjects' drawings concern the number of maxima identified in each trial and the correlation value between original and drawn topologies.

Regardless of the PMSon, subjects systematically identified the presence of the global maximum for every topology, as shown in Figure 9.4a. Secondary maxima were more frequently overlooked as their values decreased, particularly for the IOI condition. While equal in height, C_3-m_2 was more often neglected than C_2-m_1, particularly in the IOI condition. The noise level had no significant impact on the number of identified maxima.

Correlations between drawings and original topologies decreased as the topology grew more complex, significantly so between C_1 and C_2 compared to C_3 as shown in Figure 9.4b (means of $.93 \pm .04$ and $.86 \pm .08$ respectively). P and PA based explorations resulted in more precise reproductions of the original topologies compared to IOI. Interaction effect of PMSon*Topology highlight a higher dependence of IOI to the topology [$F(4, 68) = 4.47$, $p < 0.005$], particularly for the C_3 condition (related to subjects ignoring the secondary maxima as mentioned above). The noise level had no significant impact on the correlation values.

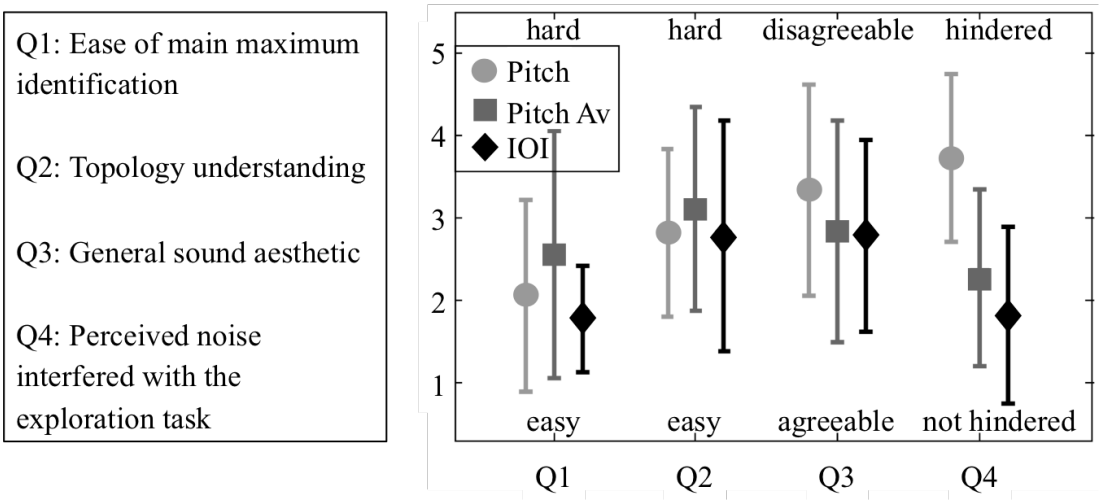


Figure 9.5: Post-experiment questions and mean and standard deviation of responses over all subjects.

The impact of PMSons and noise level on phase 2 exploration time did reflect the results of phase 1, significantly shorter with P and IOI than with PA (means of 15.1 ± 15.1 vs. 20.1 ± 18.7 sec) and significantly shorter with N_0 and N_1 compared to N_2 (means of 16.1 ± 15.7 vs. 18.2 ± 18.0 sec) . Surprisingly, the different topology conditions did not have any significant impact on the exploration time.

The topology shift had no significant impact on drawing correlation, number of maxima identified, or exploration time. Drawing correlations and exploration time of expert subjects were significantly higher than non-experts (correlation: $[F(1,16) = 4.89, p = .042]$, means of $.92 \pm .07$ vs. $.89 \pm .07$; time phase 2: $[F(1,16) = 8.11, p = .011]$, means of 26.0 ± 23.8 vs. 12.1 ± 8.0 sec for experts and non-experts respectively). Three out of the six experts identified every single maximum in their drawings, regardless of the PMSon condition. Two of these three expert subjects took on average more than twice as much time than their peers to explore the topology (means of 42.0 and 50.1 sec vs. 16.8 sec on average). In contrast, the third expert was the second fastest participant (means of 9.6 sec).

9.5.3 Subjective Evaluation

Figure 9.5 presents subject's responses to the post-experiment questionnaire, completed after each block (i.e. for each PMSon condition). The first two questions were related to phase 1 and 2 respectively. The trend suggests that subjects judged it easier to identify the main maximum and understand the topology in conditions P and IOI compared to PA, yet not significantly so. Questions 3 and 4 concerned the general

sound aesthetic and the perceived impact of the noise on the exploration task, where subjects favored IOI and PA over P, significantly for Q4 results [$F(2, 32) = 17.64$, $p < 0.001$].

When questioned, most subjects reported that the topologies explored with the IOI PMSon appeared much simpler (e.g. fewer maxima on average) than those explored with P and PA. Two of the expert subjects agreed that IOI "emphasized" higher local maxima above a certain exploration speed where the simple Pitch condition P systematically induced a mental map of the whole topology. While some subjects preferred PA over P and IOI, the PA PMSon was often criticized as "constantly limiting the maximum exploration speed" or "slow and inconsistent" as the presented information depended on the exploration speed.

9.6 Discussion

The purpose of the present study was to evaluate the capacity of the IOI and PA PMSons (detailed in Section 9.3.3) to reduce the instabilities perceived during an interactive auditory graph-based exploration of noisy data. Based on task execution time, estimation accuracy, misinterpretations, etc. the evaluation also questioned the general performance of both IOI and PA compared to the standard Pitch PMSon often used in audio-based data exploration.

Subjective evaluations indicated that both IOI and PA reduced the perception of data instabilities during the exploration, suppressing the aesthetic and fatigue issues related to Pitch-based exploration. Results otherwise showed that the impact of the noise level on subjects performance did not depend on the PMSon condition.

Quantitative results indicated that subjects handled both localization and exploration tasks (defined in Section 9.4.3) faster with P and IOI than with PA. As PA involved a running average on the data, subjects had to limit their exploration speed to perceive the topology's feature (local minima and maxima) instead of an over-smoothed version of it. Too rapid exploration speeds compared to the averaging period (see Section 9.3.3.2) are thus assumed to be responsible for the high false detection rate of PA compared to P and IOI during explorations of the most difficult topology C_3.

Fast exploration with IOI, on the other hand, was judged to emphasize the highest values of the topology compared to P that systematically induced a mental map of the whole topology. This effect of IOI on topology perception is related to the "adaptive dynamic of estimation" evoked in Section 9.2, which in essence means that the brain needs less time to evaluate tempi as the IOI decreases, and thus notices only the variations related to fast tempi during rapid explorations. The downside of this emphasis is that explorations with IOI often led subjects to discard the lowest maxima in the topology, more so in the presence of high valued ones (e.g. C_2-m_1 compared to C_3-m_2 detections in Figure 9.4a. A trend (non significant) in this figure and in Table 9.4

located in Appendix Section 9.8 suggests that explorations with P and PA were also impacted by this "masking effect". Three of the subjects identified as experts (i.e. with some musical background) easily overcame the masking effect, identifying every single maximum regardless of the PMSon condition. While this performance was accompanied by a lengthy and thorough exploration for two of them, the third expert simply reported an ease with tempo estimation due to musical training.

9.7 Conclusion

This study evaluated the performance of two PMSons designed to reduce the perception of instabilities during the real-time exploration of a noisy data stream. The two PMSons were based on Inter-Onset Interval (IOI) and Pitch Averaged (PA), i.e. a pitch mapping coupled with a time-based running average. The conducted experiment presented the audio exploration of a 1D topology using a pen tablet as the auditory viewpoint controller, where subjects had to detect the local maxima of a topology for different levels of noise applied to the data. The two PMSons were compared to a standard Pitch-based sonification in terms of aesthetics, exploration speed, and accuracy.

Results showed that IOI and PA reduced the perception of noise-related instabilities in the data compared to Pitch. The IOI PMSon furthermore preserved the system's reactivity where PA over-smoothed topologies' features for fast exploration speeds (compared to the length of the running average). IOI's downside was that it often led subjects to ignore the lowest maxima in topologies that contained high valued ones, subjects being less sensitive to slow rhythmic variations when juxtaposed with rapid ones. The performance of subjects identified as experts indicated that this masking effect could be overcome by training, after which IOI achieved Pitch-like performance and reduction of noise perception during the exploration. Said masking can be advantageously used for the selective exploration of a topology, naturally focusing listeners on a specific range of data value via the IOI sonification mapping.

A side result suggested that, for all three PMSons, the perceived amplitude of a local maximum in the auditory graph changed with the complexity of the underlying topology (further developed in Appendix Section 9.8). This observation would require further investigation through a dedicated study, as it appeared to depend on the sound parameter and may also turn out to be task-specific.

Finally, the results on Pitch subjective appreciation (low) compared to its objective performance in the exploration task (high) reinforced the findings of previous studies, suggesting that PMSons assessment will generally require both qualitative and quantitative evaluations as raw efficiency and aesthetic considerations do not necessarily point in the same direction.

9.8 Appendix: Regarding JND-based PMSon Scaling

The assumption on inter-PMSons iso-sensitivity made in Section 9.3.3 was not verified for the average subject. The variations across PMSons observed in Figure 9.4 suggest that the topologies were not perceived with the same level of detail through Pitch and IOI. This result could be seen as a consequence of the masking effect mentioned above, as the three experts who overcame this effect showed equivalent performance in terms of drawing correlation and number of identified maxima for both PMSons.

The assumption on intra-PMSon iso-sensitivity, i.e. a perceptively linear scale for all PMSons due to the JND-based design, was not verified either. Results in Table 9.4 show the average values reported for the local maxima in C_2 and C_3 in subjects drawings. While C_2 related results reasonably suggest that subjects perceived this topology without scale distortion, a comparison with C_3 evaluations indicates that the perceptive scale is distorted as the number of local maxima increases. A masking effect similar to the one described for IOI was observed, this time common to all PMSons and related to the perceived amplitude of C_2-m_1 and C_3-m_2. The fact that all three PMSons provided an equivalent accuracy on main maximum position estimation however supports an intra-PMSon iso-sensitivity for the higher topology values.

Table 9.4: Average maxima heights reported in subjects' drawings. Heights have been normalized for each drawing between maximum and minimum values drawn on the tablet, $\pm$ values give 25^{th} and 75^{th} percentiles. These results indicate a masking effect on the relative amplitude perceived as C_3-m_2, equal in height with C_2-m_1, was perceived as smaller, supposedly related to the presence of C_3-m_1.

Topology-maxID = **value**	C_2-m_1 = **0.5**	C_3-m_1 = **0.8**	C_3-m_2 = **0.5**
Pitch	**0.46** $\pm$ 0.16	**0.63** $\pm$ 0.16	**0.40** $\pm$ 0.15
Pitch Averaged	**0.46** $\pm$ 0.13	**0.61** $\pm$ 0.16	**0.38** $\pm$ 0.14
IOI	**0.35** $\pm$ 0.15	**0.57** $\pm$ 0.17	**0.27** $\pm$ 0.12

A subsequent dedicated study would be necessary to clearly assess both iso-sensitivity assumptions for the scaling functions presented in Section 9.3.3.

Part III

Prototyping and Evaluation

The work presented here concerns research carried out related to the prototyping of an audio direction finder. A virtual prototype implementation is suggested, based on the synchronization of a raytracing propagation model to the elements within a virtual scene. Said prototype is employed to evaluate the behavior and performance of several direction finder designs during simulated search tasks in the virtual scene. A physical prototype is implemented to validate the results issued from the simulation and further assess the performance of audio designs compared to other victim localization techniques during lifelike search and rescue operations.

Chapter 10

Virtual Prototype

Contents

Until recently, human-device interface assessments primarily involved a lengthy prototyping process. Engineers and researchers usually had to construct real prototypes to gather some relevant feedback on their design from ergonomists and potential end-users. Recent developments in Virtual Environment (VE) related technologies have given rise to the experimental process known as "Virtual Prototyping", defined as *"A computer simulation of a physical product that can be [...] tested from concerned product life-cycle aspects [...]"* by Wang in [202]. Applications range from surgery training, e.g. using haptic feedback [203] sometimes combined with visual rendering based on head-mounted display [204], to car design evaluations on social network platforms [205]. The process has considerably improved ergonomic studies, allowing for low-cost and fast design evaluation in supervised VE [206].

10.1 Introduction

This chapter details the implementation of a Virtual Prototype (VP) of the different radio Direction Finder (DF) designs presented in Part I. The VP core architecture was the same for every DF designs, composed of a virtual scene in which users may navigate using a virtual DF to find their way towards a virtual Radio Frequency (RF) emitter.

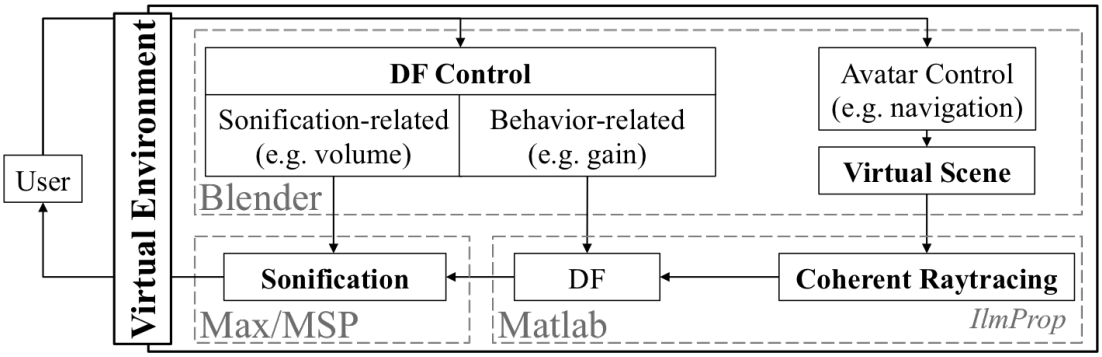


Figure 10.1: Simplified architecture of the Virtual Prototype.

The objective of the VP implementation was to allow the evaluation of dynamic performance and behavior of the DF designs during a the navigation, as opposed to the static evaluations proposed in Chapter 4.

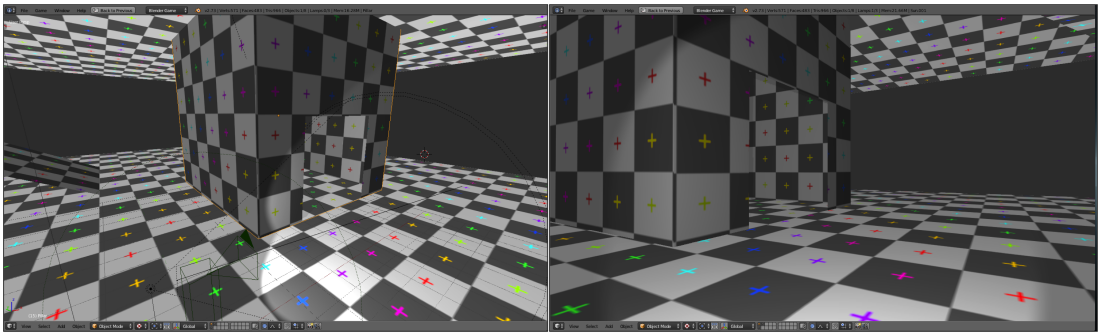
Both emitter and DFs were simulated based on the Ilmprop propagation model coherently coupled with the virtual scene. "Coherent" here indicates that the propagation model took into account the different elements that composed the virtual scene (walls, obstacles, etc.) to simulate RF propagation.

A short VP-based study of different DF designs performance in simulated search conditions is presented at the end of this chapter, Section 10.2. Next Section details the VP implementation, from the creation of the virtual scene to the real-time DF-based navigation. Ultimately, the VP will be used in parallel with the physical prototype introduced in Chapter 11 to evaluate and improve DF designs after the assessment of its ecological validity detailed in Chapter 12.

10.2 Virtual Prototype Implementation

In what follows, the term Virtual Prototype refers to the virtual DF and its simulated behavior in the VE. Said VE is composed of the VP, a virtual scene coupled with a raytracing propagation model, and a tangible interface to control the DF and navigate in the scene. The virtual scene is designed in Blender [207], a 3D content creation software which also handled navigation and DF-related interactions. The RF propagation model is based on IlmProp [71], the Matlab RF raytracing propagation model introduced in Chapter 4. As IlmProp is a 2D raytracing model, the VP is limited to search scenarios in an horizontal plane. The DF output sonification is processed via Max/MSP, the sound processing software used for the sonification design in Chapter 8. Communications between Blender, Matlab, and Max/MSP are handled through the Open Sound Control[1] (OSC) protocol. Figure 10.1 details the overall VP architecture.

[1]Inter-software communication protocol based on the User Datagram Protocol (UDP).



(a) (b)

Figure 10.2: Virtual scene in Blender. (a) Edition of the virtual scene, move box and camera (bottom of the picture) are non-textured objects, not visible during the rendering of the virtual scene. (b) Rendering of the virtual scene, first-person view.

The VP was initially implemented on a laptop. Section 10.2.3 details the adaptation of the VP to a CAVE (Cave Augmented Virtual Environment) architecture to simulate "lifelike" rescue operations. While on the initial implementation interactions were limited to keyboard and mouse controls, the final CAVE implementation enables more realistic navigation and interaction paradigms.

10.2.1 Virtual Scene

Much like a first-person video game, the virtual scene was composed of a virtual world and an avatar to represent the user's viewpoint, along with a set of rules that defined the potential interactions between the two. As the objective of the VP was to assess the DF in realistic conditions, several virtual worlds were designed to reproduce typical search environments (forest, urban areas, buildings, etc.). The user navigated through the virtual world from a first-person perspective, which reduced the avatar's implementation to a camera attached to a "move box", used to *physicalize* its movements (collisions with walls rather than going through them). In the laptop-based VP implementation, the user controlled their position and viewpoint in the virtual scene via keyboard and mouse controllers. The targeted cellphone was represented by a basic mesh added on the floor of the virtual scene. Figure 10.2 illustrates a basic virtual scene in Blender.

10.2.2 RF Propagation in the Virtual Scene

Simulating a coherent RF propagation in the virtual scene required the use of said virtual scene as an input for the raytracing model geometry, not unlike the implementation proposed in [208]. The geometry of the elements that composed the scene

Table 10.1: Example of IR dataset size and computation time using the IlmProp raytracing model for various grid steps in the imported scene of Figure 10.3b, defining the DF antenna array as a 8 elements UCA. For reference, the grid step illustrated in Figure 10.3b and through the video examples of Section 10.3 is 2 m. For the walking speed defined in the virtual scene, this corresponds to IR updated approx. every second for a user walking in a straight line. (Calculations executed on a 2.3 GHz Intel Core i7 macbook pro with 4Go of RAM at 1600 MHz running Matlab R2012a).

Grid step size	Computation time	IR array size (.mat)
10 m	15 sec	1.6 Mo
4 m	1 min	5.2 Mo
2 m	5 min	25 Mo

was imported in the raytracing model and their nature regarding RF propagation was defined (refraction, diffraction, scattering coefficient, etc.). User's avatar and target meshes were replaced by their RF equivalent during the importation, i.e. a DF antenna array for the user (receiver) and a cellphone-like omnidirectional antenna for the target (emitter). From this setup, a typical run of the IlmProp raytracing model produced an array of Impulse Responses (IRs) representing the different RF paths from emitter to receiver's antennas (an IR for each of the receiver's antennas).

As the objective of the VP was to navigate in the virtual scene with the DF to find a fixed RF emitter in real-time, the IRs for potential positions of the DF (user) in the explorable space were calculated and stored prior to the navigation. To limit the size of the resulting IR dataset, a grid of possible positions was predefined based on the virtual scene size and a grid step. As the user navigated in the virtual scene, their position was approximated as the nearest grid point and sent to a Matlab script that selected the associated IR array and ran the DF-related estimation.

The grid step was defined as a tradeoff between the IR dataset computation time/size and the perceived continuity of the DF estimations during the navigation. Too large a grid step resulted in scarce IR index updates, i.e. whole areas of the virtual scene that corresponded to the same IR array producing a unique Direction Of Arrival (DOA) estimation. On the other hand, the smaller the grid step, the longer the IlmProp offline calculation and the larger the produced IR dataset.

Figure 10.3 illustrates the overall pipeline employed to create contextualized RF propagation in the virtual scene. Table 10.1 reports typical IlmProp computation time and stored IR dataset size for various grid step values applied on the virtual scene of Figure 10.3b.

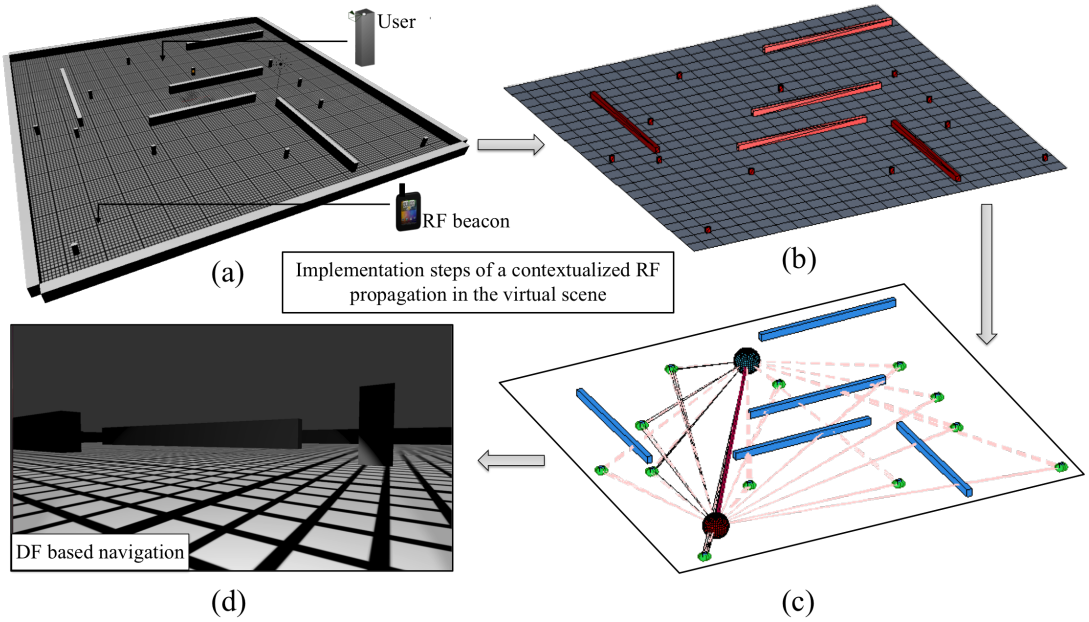
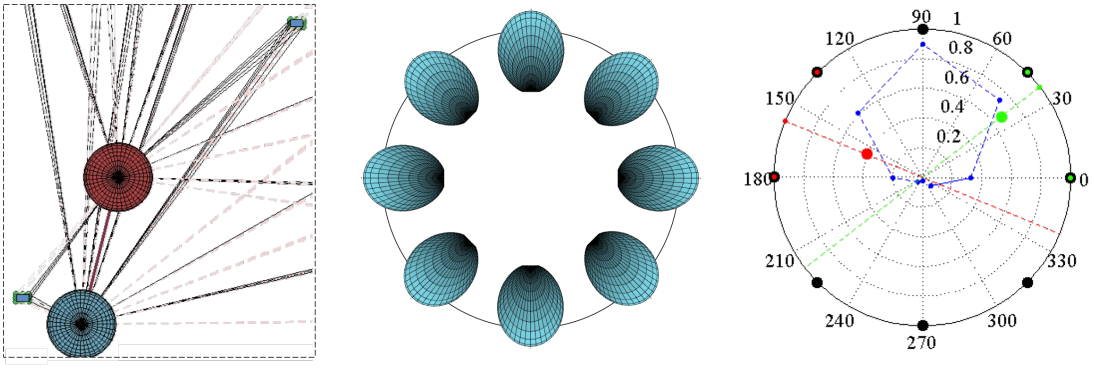


Figure 10.3: Detail of the different steps involved in the simulation of contextualized RF propagation in the virtual scene for DF based navigation. (a) Creation of the virtual scene in Blender. Positioning of both User and target (RF beacon) in the scene. (b) Import of the scene geometry in Ilmprop, definition of virtual objects behavior regarding RF propagation. Definition of a grid and its associated grid step for user's position approximation and IR index selection (pictured grid is 25×25 m^2 for a 2 m grid step). (c) Computation of the IRs from emitter (RF beacon) to receiver (virtual DF co-localized with the user) for each potential position of the emitter on the grid (grid nodes). For this illustrative implementation, scatterers (green dots) were defined around every pillar-like objects in the scene, reflecting the RF signal to simulate multi-path propagation. Walls were made to simply block the RF paths (dotted pink lines). (d) Screenshot of the user's viewpoint during the navigation in the virtual scene.

The IR dataset pre-processing assumed a constant DF orientation[2] for each potential DF position. This assumption was made to avoid a massive increase in the size of the pre-processed IR dataset, i.e. processing $N_{\text{position}} \times N_{\text{orientation}}$ IR arrays instead of the N_{position} processed so far. Rather, DF's orientation was "manually" added to the current IR array during the navigation, differently so according to the considered DF design:

For automatic DF designs, i.e. whose output consisted of an estimated DOA or spatial power spectrum (see Chapter 3), the current DF orientation was added to

[2]DF orientation concerned the azimuthal plane only because of the 2D constraint on the raytracing model.



(a) (b) (c)

Figure 10.4: Creation of the IR of two directional antennas based on the IR of an UCA. (a) RF propagation scenario, the red sphere is the emitter, the blue is the UCA receiver. (b) UCA geometry composed of 8 gaussian antenna of individual apertures 68° in azimuth. (c) Calculation of the received signal strength (red and green dots) of a fictive pair directional antennas for two different orientations (red and green dotted lines at 38° and 158°) through a summed average over the two nearest neighboring UCA elements (2×2 red and green dots circled in black on circle's circumference). This synthesized directional antenna had an aperture of 60° (i.e. the linear gain at ±30° is half the gain at 0°).

the estimated DOA or applied as a circular shift to the spatial power spectrum. To produce results similar to the exhaustive calculation of $N_{\text{position}} \times N_{\text{orientation}}$ IR arrays, this method required that the measured IR did not depend on the DF (antenna array) orientation. This need for rotational invariance in the azimuthal plane led us to select an Uniform Circular Array (UCA) as a basis for the DF designs assessed via the VP. This assumption made on the antenna array geometry limits the potential of the VP in terms of DOA estimator assessment, yet UCAs have the advantage of being representative geometries for assessing the behavior of most DOA estimators. Tests showed that the estimated power spectrum complied with the rotational invariance hypothesis for UCA arrays composed of $M \geqslant 6$ omnidirectional antennas.

For manual DF designs, composed of one or more directional antennas, a likewise M-element UCA was used for IR arrays pre-processing, this time composed of M directional antennas. The Received Signal Strength (RSS) of "virtual" DF's antennas were simulated based on a weighted average of the RSS of the two UCA antennas closest to the current orientation of the considered virtual antenna. As for automatic DF designs, the UCA geometry was selected to because of its rotational invariance. Modifying UCA's number of antennas M and their directivity allowed one to simulate various directional antenna *apertures* (i.e. directivity). Figure 10.4 illustrates the use

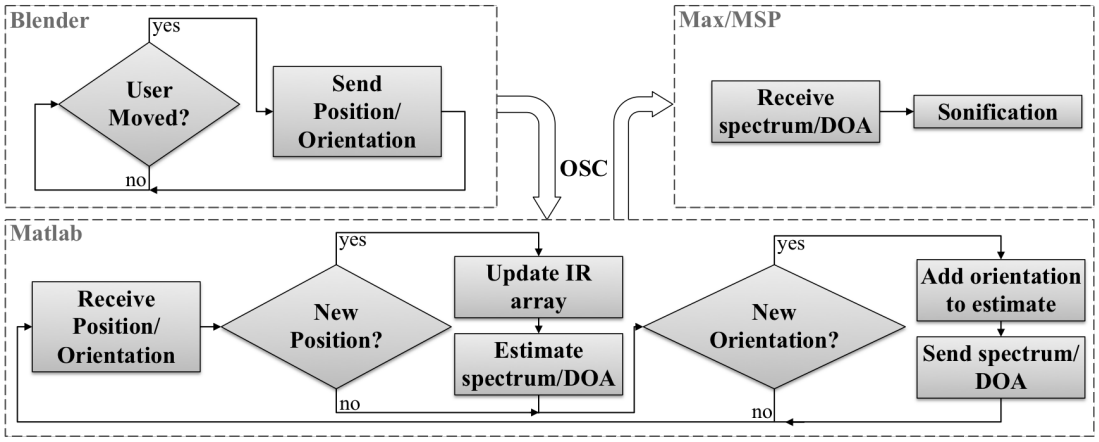


Figure 10.5: Simplified sequencer of a run of the VP.

of an UCA composed of $M = 8$ gaussian antennas to simulate a pair of directional antennas with a 60° aperture (similar to the one used for the physical prototype introduced in Chapter 11).

Figure 10.5 details the steps involved in a typical run of the VP once the overall preprocessing achieved. Videos //assets/partIII/2.* virtual-prototype-v*-search.mp4 illustrate DF-based navigation in the virtual scene, further introduced in Section 10.3.

10.2.3 Lifelike Rescue Operations, CAVE-Based Virtual Prototyping

The overall VP implementation has been ported to the EVE platform [209], a CAVE architecture, to simulate lifelike rescue operations. EVE employs active adaptive stereoscopic rendering on four rear-projected screens coupled to a cluster of seven cinema-sized projectors each controlled by an i7 computer, altogether achieving about 36 m^2 of high-definition projection space. The VE was rendered through BlenderVR[3] [3], an extension of the Blender software into a complete authoring tool for VR real-time interactive scene rendering. The sonification was still handled by MaxMSP coupled with an RF wireless headphone module for mono, stereo, and binaural rendering. Users explored the VE from a fixed position allowing for an approximate 180°h×110°v field of view, observed through tracked stereoscopic shutter glasses (c.f. Figure 10.6). To provide some ecological equivalent of walking, navigation in the VE was based on the combination of a Nintendo Wii Balance Board and Nunchuk (single joystick along with analog buttons) through a walk-in-place metaphor. The walk-in-place implementation was a simplified version of the algorithm introduced in [210], where the Nunchuk joystick controlled Z axis rotations of the user's move box (and attached viewpoint) within

[3] `http://blendervr.limsi.fr/doku.php`. BlenderVR developments are detailed in Appendix A.

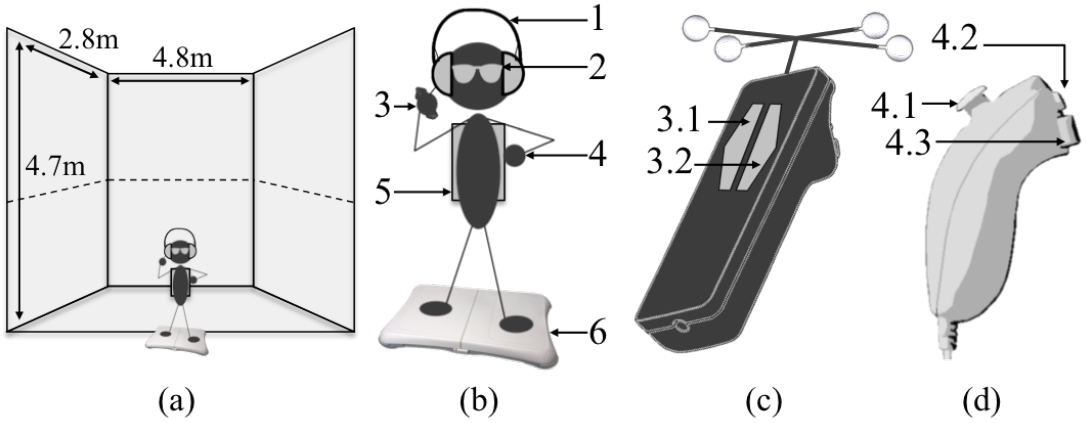


Figure 10.6: Overview of the CAVE-based Virtual Prototype setup. (a) EVE system. Only the lower half (below the dotted line) of the vertical projection screens was used for the rendering (using two projectors per front screen for higher video quality). (b) Hardware used to interact with the virtual environment. (b1) Headset and its associated RF receiver. (b2) Tracked stereoscopic shutter glasses. (b3) 6 DoF tracker and its associated analog buttons, used to interact with the virtual DF (e.g. to aim the directional antenna). (b4) Nintendo Nunchuk. (b5) Loaded backpack, used to reproduce the physical load related to the DF. (b6) Nintendo Wii Balance Board to create the walk-in-place metaphor. (c) Detail of the 6 DoF tracker, its two analog buttons (c3.1 and c3.2) used to map DF related interactions. (d) Detail of the Nintendo Nunchuk. (d4.1) Joystick used to control Z axis related rotation of the virtual vehicle (i.e. viewport). (d4.2 and d4.3) analog buttons used for navigation related interactions.

the VE, as EVE is not a full 360° system. A 6 DoF (Degree Of Freedom) tracker coupled with some analog buttons was used for DF-related interactions. The video //assets/partIII/1.1 virtual-prototype-cave-illustration.mov presents an illustration of the CAVE VP for a simple DF design based on a hand-held directional antenna.

The main asset of the CAVE compared to a laptop implementation was that users were no longer in front of a "video game". No longer constrained to use keyboard interactions, users were more willingly immersed in the 3D environment and could completely focus on the search task. The inclusion of an ambisonic audio system integrated in the EVE architecture allowed for the simulation of realistic soundscapes, e.g. to assess the masking effect of potential environmental sounds onto sonification designs. The walk-in-place metaphor, supposing walk-like movements, allowed the evaluation of fatigue related to the wearing of different DF designs during the task. This CAVE-based implementation was used to assess the ecological validity of the VP, see Chapter 12.

Table 10.2: Description of the three DF designs implemented as VP for the informal experiment. All three designs were assumed to be mounted on user's helmet, i.e. with the same orientation as user's head (or equivalently, user's viewpoint in the VE).

DF_1	1 directional antenna, sends RSS measurement to a Geiger counter sonification. Positive mapping: the perceived rhythm increased with the RSS. The antenna orientation is aligned with user's viewpoint.
DF_2	8-element UCA array, sends MUSIC-based DOA estimation to binaural spatialization. The considered UCA is the one illustrated in Figure 10.4b. A geiger counter metaphor is overlaid to the binaural rendering, based on the average RSS received by UCA's antennas.
DF_3	2 directional antennas, sends the 2 RSS measurements to geiger counter overlaid to binaural spatialization based on RSS ratio. The orientation of both antennas is fixed at $\pm 45^\circ$ compared to user's viewpoint. The geiger counter is ruled by the averaged value over the two RSS. If both antennas measure the same RSS, the sound is heard in front. The sound progressively shifts to the right as the ratio RSS_{right}/RSS_{left} increases compared to 1 and reciprocally to the left as it approaches 0.

10.3 DF Performance Evaluation Based on the Virtual Prototype

This section presents the observations issued from informal tests with the DF VP, comparing the performance and behaviors of three different audio DF designs. This comparison is followed by a general discussion of foreseen issues related to the design of an audio DF for search and rescue operations.

10.3.1 Method and Experimental Design

The three DF designs considered are described in Table 10.2. All three were implemented as VPs on a laptop, following the architecture described in Section 10.2. DF_1 simulated a typical manual DF, mapping the RSS of a directional antenna to a Geiger counter sonification. DF_1 was a typical implementation of the "Geiger counter for cellphone signals" discussed in Chapter 9[4]. A control over the RSS sensitivity gain was added prior to the sonification, allowing users to manually compensate for the large variation dynamic of the RSS evoked in Section 8.5, discussed in more details Chapter 12. There were 5 gain levels, coupled with a spearcon-based audio feedback

[4]Without any noise in the measured RSS though, as the IRs processed prior to the navigation.

when gain level was modified by the user ("five", "four", etc. heard at the beginning of //assets/partIII/2.1 virtual-prototype-v1-search-1Rx.mp4).

DF_2 simulated a typical automatic DF, mapping an estimated DOA to the spatial position of a sound source. DF_2 was based on binaural rendering coupled with a MUSIC estimator using an 8-element UCA array, similar to the one introduced in Section 4.3.3. The objective was to perceive the sound as if coming from the estimated DOA. A Geiger counter metaphor was added based on the average RSS of the UCA elements, to provide users with some feedback on their distance relative to the RF target. As with DF_1, having a rapid feedback with regards to changes in RSS which related to distance allowed users to determine whether they were approaching or distancing themselves from the target. The same spearcon-based control over the sensitivity gain than for DF_1 was provided.

DF_3 was proposed as a tradeoff between DF_1's simplicity and DF_2 foreseen efficiency. Based on a pair of directional antennas, DF_3 implemented a Geiger counter metaphor coupled with a binaural rendering weighted by the antennas' RSS ratio. Due to the antennas specific orientation, the sound spatialization indicated which direction the user needed to turn the DF in order to increase the average RSS.

Five subjects participated in the experiment. Amongst the subjects were two potential end-users from the SDIS[5] (french equivalent of USAR[6] firefighters). For each DF design, subjects had to search for 4 targets in the virtual scene illustrated in Figure 10.3a guided by the sonification metaphors. After a brief training where they were introduced with VP controls and sonification paradigms, subjects started the first search. The initial position of both User (subject viewpoint) and RF beacon (target) in the virtual environment were randomized, avoiding trivial scenarios were subjects would start the search in front of the target. When they reached the target, subjects were repositioned in the virtual scene and notified to search for the next target. Targets were represented by a cellphone-sized object, see Chapter 12 for further illustration.

The results discussed below are not issued from formal measurements, but rather from subjective assessment of participants performance by an external examinator. These search tests were intended as a pre-evaluation of the different audio DF designs performance. As a future study, this experimental design could be extended (more subjects, quantitative logs, etc.) to allow rigorous comparison of the considered audio DF designs.

[5]Service Départemental d'Incendie et de Secours.
[6]Urban Search And Rescue

10.3.2 Results and Designs Comparison

The video //assets/partIII/2.1 virtual-prototype-v1-search-1Rx.mp4 illustrates a typical use of the design DF_1. While the search in the video was deliberately extended for the sake of illustration, DF_1 offered reasonable performance for a minimum design complexity. Its main inconvenient was that it forced subjects to regularly perform 360° scans, interrupting their navigation to estimate the DOA of the maximum RSS at a given location in the exploration space.

A typical use of DF_2 is illustrated in video //assets/partIII/2.2 virtual-prototype-v2-search-UCA-DOA.mp4. Compared to DF_1, the DOA estimation slightly increased the search speed. The binaural rendering made the target's direction immediately identifiable, avoiding the systematic 360° scan to find the maximum RSS. Compared to the results concerning the automatic search of Chapter 5, errors on DOA estimation related to multipath propagation did not present an issue with the VP. The average DF user was able to adapt the search pattern and in a sense to "integrate" information during the navigation, making decisions based on subjective weighted averages over several estimations, compensating for scenarios where the DF produced an erroneous DOA estimation. The only drawback of DF_2 was its implementation complexity, mainly related to the number of antennas that composed its array[7].

A search with DF_3 is illustrated in video //assets/partIII/2.3 virtual-prototype-v3-search-2Rx.mp4. Without the 8-element complexity of DF_2, this design still allowed users to avoid the need to perform the 360° scans of DF_1 by adding a simple dynamic left/right information.

10.3.3 Discussion

Implemented as a VP, DF_3 seemed a promising design as an in-between solution between implementation complexity and navigation performance.

Based on the observation of users performance during the search, and on users' comments following the experiment, the VP highlighted potential issues with the foreseen DF prototype. All three DF designs were subject to the dynamic of variations of RSS, following an inverse square law regarding the distance from DF to RF target in the best case scenario (free field hypothesis). Most subjects would have expected a linear decrease of the perceived distance, e.g. 1 m in the correct direction should always "sound the same". Accordingly, the inverse square law dynamic forced to question the final variation range of the sound parameter. For a constant RSS sensitivity gain or a careless design of the RSS sensitivity gain levels, dynamic resulted in large variations of the sound parameter when hovering close to the target and nearly imperceptible for large distances.

[7] As discussed in Chapter 6, the number of antennas usually has a severe impact on the size of the array.

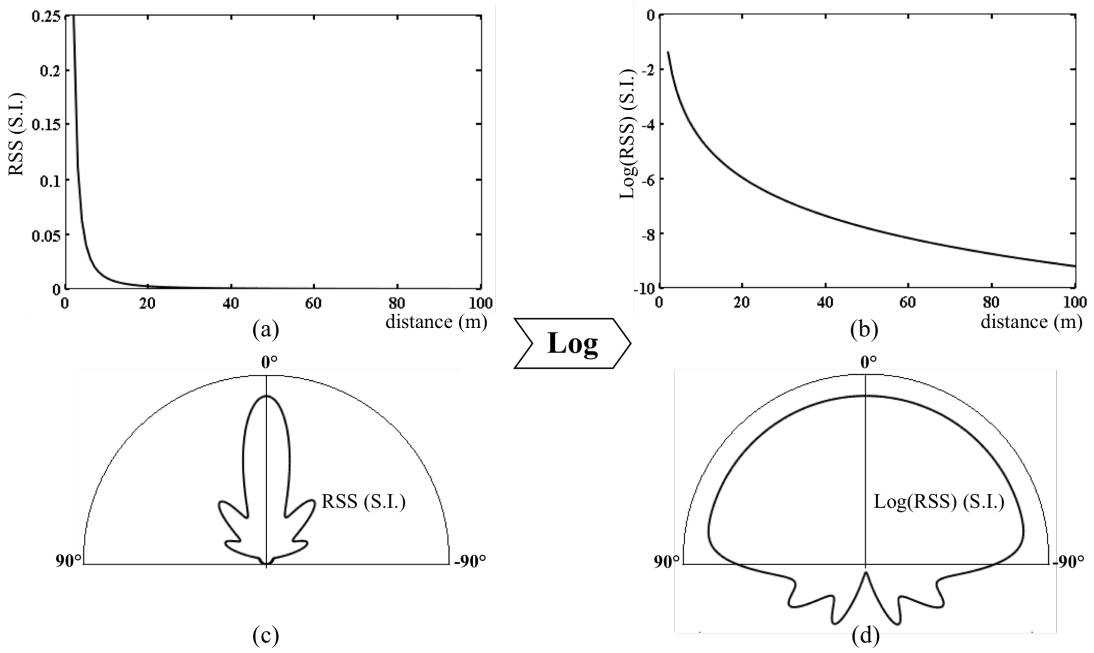


Figure 10.7: Illustration of the impact of a logarithmic mapping function applied to the RSS of a directional antenna. (a) Original 1/distance^2 RSS dynamic. (b) "Linearization" of the RSS dynamic based on a logarithmic function. (c) Original directional antenna beam-pattern (typical yagi antenna). (d) Impact of the linearization of (b) on the antenna beam-pattern. While the mapping from (a) to (b) simplifies the sonification mapping, its impact on the beam-pattern transforms the directional antenna of (c) into a cardiod antenna (d) with scarcely more directivity than an omnidirectional antenna.

Aside from the sensitivity gain levels, a solution could be to use some logarithmic scaling function prior to the sonification, to provide a linear increase of the perceived sound parameter as the target draws nearer. Such a scaling function would however flatten the beam-pattern of the directional antennas in DF_1 and DF_3 as illustrated in Figure 10.7, drastically diminishing their directionality, essential during the search. Another solution considered was to dynamically normalize the RSS (through the sensitivity gain), manually triggering said normalization to adapt it to a given DF-to-target distance. Dynamic normalization however led to a complete loss of distance-to-target perception if not coupled with e.g. a PMSon over the current gain level, which simply shifted the inverse square law dynamic problem.

Finally, the gain for each level could be adjusted so as to provide a roughly piecewise linearization of the global inverse square law. This process is employed and detailed in the DF implementation in Chapter 12. Personal observations suggest that a high performing functional solution would involve a semi-automated gain control (to en-

sure users' awareness of their progression) coupled with a sonification of the gain level overlaid on the main auditory stream. The spearcon sonification feedback of the RSS sensitivity gain implemented for all three DF_i is one option amongst others, e.g. replaced in the implementation of Chapter 12 by a change in the main Geiger counter sonification timbre as the sensitivity gain is modified.

An advantage of the VP over classic prototyping is its malleability, which allowed us to quickly assess potential designs while providing complete control over the search complexity. An informal exploration was for example conducted to compare the visual and audio modality for the three DF designs mentioned above. Said test simply required the implementation of a visualization widget with either a compass for DF_2 or a RSS level meter for DF_1 and DF_2[8]. No significant difference was observed between both modalities while the audio feedback naturally allowed users to completely focus on the search environment.

Another informal exploration was conducted where searches in the virtual scene were performed using the true target direction as an input for the sonification, without either propagation nor estimation. This allowed further assessment of the potential of audio against visual modality for the navigation aid aspect of the DF, especially in 3D environments (multi-storey parkings, buildings, etc. since the 2D constraint imposed by the propagation model was no longer present). Both modalities produced roughly equivalent performance. As the environment grew more complex (doors, floors, dead ends, etc.) the knowledge of the true direction sometimes led to tedious searches where the "nearest path" would have been more helpful than a straight line through obstacles. As put forward in [19], a DF that indicates the nearest path towards the victim[9] is probably preferable over a system encouragin users to walk through concrete walls.

Finally, the VP allowed one to investigate potential masking effects of environmental sounds on the DF sonification in a controlled environment. Using the ambisonic system in the EVE architecture or built-in laptop loudspeakers, we were able to assess the intelligibility of sonification designs in noisy conditions, e.g. with audio files such as the //assets/partIII/2.4 street, alarms.mp3. It appeared that environmental sounds should not be an issue as long as users have some control over the system volume and the "harshness" of the sounds involved in the sonification. An adjustable comb filter can e.g. be used to control the latter.

Bonephones proved useful in situations were users needed to maintain some level of auditory attention to monitor their surroundings or exchange information with other individuals during the search. Bonephones should not however be used in very noisy environments, as it becomes hardly possible to focus on the sonification, particularly

[8] Two level meters were actually required for DF_2, one for average RSS, the other for RSS ratio.

[9] " *[...] similar to what rescue dogs would find because of ventilation.*", from [19].

since the dynamic range of bonephones is limited [211]. For excessive input level, the "audio" output of the average bonephone will indeed be perceived as actual vibration on the temples[10] (a feature that could however be used for mixed audio-tactile feedback though).

10.4 Conclusion

This chapter detailed the implementation of the DF Virtual Prototype (VP). Said implementation consisted in the creation of a Virtual Environment (VE) composed of a virtual scene, a user avatar, a virtual DF, and a virtual RF target. RF propagation from the target to the DF was simulated based on the virtual scene geometry, constrained to 2D propagation scenarios because of the simple raytracing model employed. As the objective of the VP was to assess the behavior of DF designs during real-time navigations, most of the calculations related to the RF propagation were executed offline, storing the IRs for a uniform grid of discrete DF positions in the virtual scene. The overall VE was adapted to a CAVE architecture to further enhance its ecological validity, further discussed in Chapter 12.

Informal tests and search simulations in the VE allowed to compare the performance of three different DF designs. Observations suggested that the performance of a manual DF design could match those of an automatic DF design (based on automatic DOA estimation), providing the use of several directional antennas to increase the amount of sonified information and assist users in their search for a local RSS maximum. Finally, the VP allowed us to anticipate potential issues related to the implementation of a physical prototype. Scaling issues are e.g. to be expected, related to the non-linear decrease of the RSS as the DF-to-target distance decreases.

[10] In addition to saturation and distortion of the auditory signal.

Chapter 11

Physical Prototype

Contents

This chapter details the implementation of a Physical Prototype (PP) of the audio Direction Finder (DF). Based on Software Defined Radio (SDR), said implementation allowed us to quickly prototype various DF designs that reused the sonifications developed through the Virtual Prototype (VP).

After a brief introduction to the assets of SDR and a description of the PP implementation, this chapter presents the results of a preliminary study aimed at assessing the performance of DF designs DF_1 and DF_3 introduced as VP in Chapter 10 in real search conditions. The study was conducted in collaboration with firefighters from the SDIS[1] (French equivalent of Urban Search And Rescue squads) and supervised by an ergonomist intern involved in the research project, in charge of the assessment of audio DF design ergonomics. Other victim location solutions (avalanche transceiver, trained dogs, etc.) were deployed during the search simulations involved in the study, allowing to assess the performance of the audio DF relative to other state-of-the-art systems.

[1] Service Départemental d'Incendie et de Secours.

11.1 Introduction: The Assets of Software Defined Radio

Software Defined Radio [212, 213] typically refers to systems composed of a Radio Frequency (RF) front-end connected to a computer which handles the majority of signal processing routines. Due to increasing personal computer power, SDR architectures have become the fastest way to design and improve RF prototypes. Functions such as filters, modulators/demodulators, detectors, or coding schemes, can now be implemented, tested, and improved without any hardware intervention, using high-level programming languages (C++, Python, etc.).

Initially referred to as "*Software Radio*" [214], SDR is closely related to the notion of cognitive radio [215]. Both concepts emerged as governments and industries expressed the need for more flexible communication systems in the early 1980's regarding architecture reconfigurability and RF bandwidth optimization. As the required technologies have matured, SDR is now commonly used to develop digital RF mockups [216], prototype wireless standards [217, 218], or design cost-effective solutions for niche markets [219, 220].

Governments and industries now perceive the immense potential of SDR as a threat as well as an opportunity [221]. Offering instantaneous access to any RF waveform, combined with a continuous increase in personal computer performance, it is commonly acknowledged that SDR will force authorities to review some of the telecommunication security standards. The SDR-based air traffic communication monitoring presented in [222] or the wireless protocol attacks described in [223] are but perfect illustrations of the risks inherent to the democratization of the access to RF transmissions.

Except for business-to-business solutions, the most widely spread RF front-ends for SDR applications are currently the series known as Universal Software Radio Peripherals (USRPs), first developed by Ettus Research in 2005 [224]. Some efficient SDR chipsets were also developed within the amateur radio community [225] or adapted from commercial transmitters [226]. On the software side, Matlab Simulink [227] and Labview [228] have adapted their framework to propose ready-to-use platforms, already widespread in academical applications [229, 230]. Open source SDR software like Click [231] or GNU Radio [232, 233] also emerged, the latter being now readily accepted as a reference by the SDR community [234, 235].

Several studies have previously implemented SDR-based DFs, processing Direction Of Arrival (DOA) estimation through Beamforming [236], Pseudo-Doppler [237], or MUSIC algorithms [238].

The PP implementation, both hardware and software, is detailed Section 11.2. Section 11.3 presents a preliminary study on the evaluation of DF_1 and DF_3 designs introduced in Chapter 10, implemented as PPs.

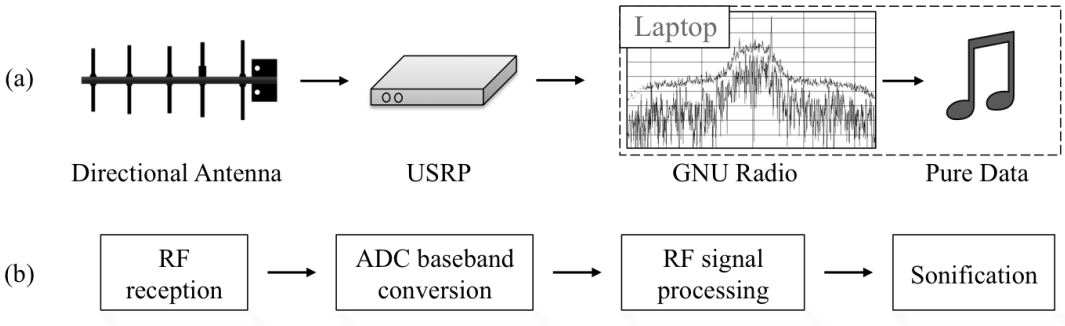


Figure 11.1: General architecture of the Physical Prototype. PP elements (top) and block diagram (bottom).

11.2 Physical Prototype Implementation

This section presents the implementation of the physical DF prototype developed to support the investigations on sonification design. The PP implementation was based on both USRP hardware and GNU Radio software to receive and process RF signals. Processed RF signals were then sent to Pure Data, an open source audio processing software [239], for sonification. In initial versions, Matlab Simulink Universal Hardware Driver (UHD) was used instead of GNU Radio, but was finally discarded in order to focus on a license-free prototype based on a Linux architecture because of industry-related requirements. The software/hardware implementation described herein is based on the "RF Geiger counter" referred to as DF_1 in Chapter 10, where the Root Mean Square (RMS) of the signal received by the directional antenna was mapped to the repetition period of audio bursts. DF_3-specific design details are reported Section 11.2.4.

11.2.1 Hardware

The general architecture of the PP implementation is detailed in Figure 11.1. The choice was made to limit the PP implementation to manual DF designs based on a maximum of 2 antennas, as opposed to a prototype allowing the test of automatic DF designs, requiring antenna arrays composed of a minimum of four to five antennas [81]. This choice allowed to considerably reduce the PP complexity and the time required for its implementation, while obviously limiting the range of its application in terms of DF designs assessment.

The RF front-end was composed of a dual-band directional antenna and an USRP. The directional antenna selected was the Yagi-Uda antenna illustrated in Figure 11.2, designed to operate in both GSM900 and DCS1800 bands (see Table 2.2). The USRP used was a N210 coupled with a WBX daughterboard, basically a FPGA coupled with a RF transceiver. The FPGA assumed the conversion of RF analog data received by

Band	GSM 900 (806-960 MHz) GSM1800 + UMTS (1710-2500 MHz)
Gain	GSM900: 8.2 dBi GSM1800/UMTS: 9.5 dBi
Aperture V/H	GSM900: 60°/90° GSM1800/UMTS: 55°/70°
Polarization	Vertical
Dimensions H/W/L	293 x 210 x 65 mm
Weight	0.94 kg

Figure 11.2: Illustration and specifications of the Yagi-Uda directional antenna used for the DF prototype. Its radiation pattern is a cardioid defined by Apertures V and H. For example, according to the figures of Aperture GSM900 H, the antenna gain for a GSM900 source emitting above the pictured antenna would be of 8.2 dBi and would be decreased by 3 dB if the source was shifted left or right from $90°/2 = 45°$.

the antenna into baseband digital data, sent to the host computer through an ethernet cable. Figures 11.3 and 11.4 detail URSP's specifications and packaging.

11.2.2 Software

The software was hosted on a computer connected to the USRP. A Python script based on the GNU Radio library (Python and wrapped C++ libraries) processed the RF data sent from the USRP. The same Python script forwarded these processed data to Pure Data via UDP for sonification. As GNU Radio and Pure Data UDP formatting mismatched, a Python server or "bridge" was used to transform UDP packets sent by GNU Radio into OSC messages forwarded to Pure Data. Figure 11.5 is a reminder of DF_1's principle while Figure 11.6 illustrates both GNU Radio and Pure Data PP interfaces.

GNU Radio

Using GNU Radio related Python library, or the user friendly block-based editor GNU Radio Companion (GRC), the implementation of the RF chain primarily consisted in the manipulation of predefined functions (blocks). Each block was parametric, enabling simple modification of central frequency, bandwidth, gain, etc. For DF_1, the signal processing was achieved with four blocks that calculated the RMS of the received signal and sent it through UDP to the UDP-to-OSC bridge. A 100 kHz low-pass filter centered around the RF frequency of the emitter (described in Section 11.2.5) was added upstream of the RMS calculation to limit potential interferences from other RF emitters. As seen in Figure 11.6, GRC also provided some useful GUI (Graphic User Interface) elements used to monitor and modify the processing chain in real-time.

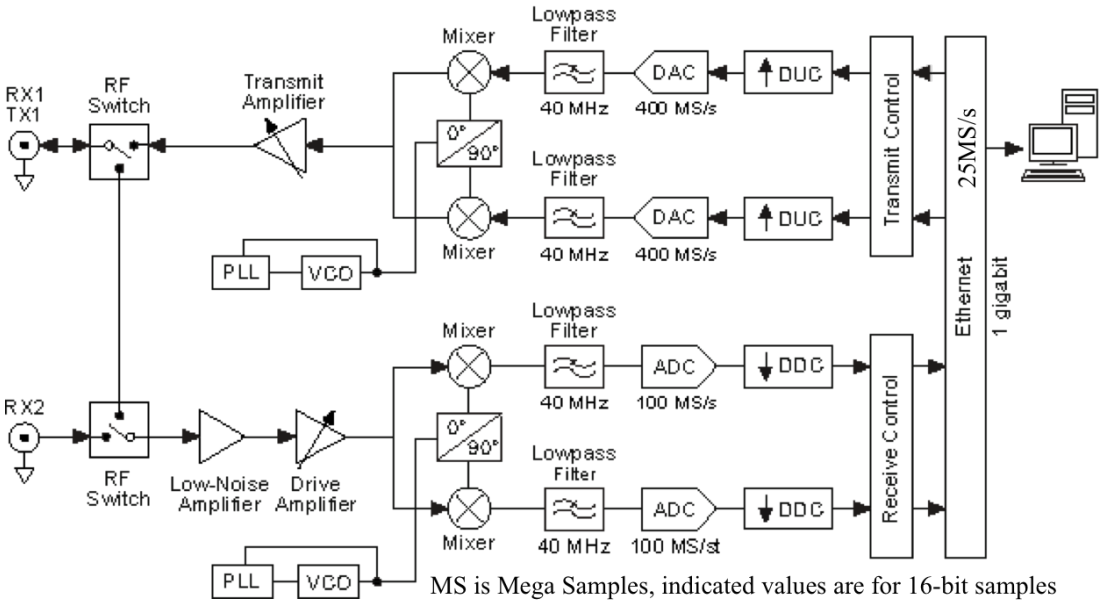


Figure 11.3: Block diagram of the USRP N210 with its WBX daughterboard. In this configuration, the USRP can receive signals from 50 MHz to 2.2 GHz with approx. 20 MHz of usable bandwidth (the original 40 MHz analog bandwidth of the WBX is limited by the GigE interface to the host computer) and a noise figure of 5 dB. Figure taken from http://www.ettus.com/kb/detail/usrp-bandwidth.

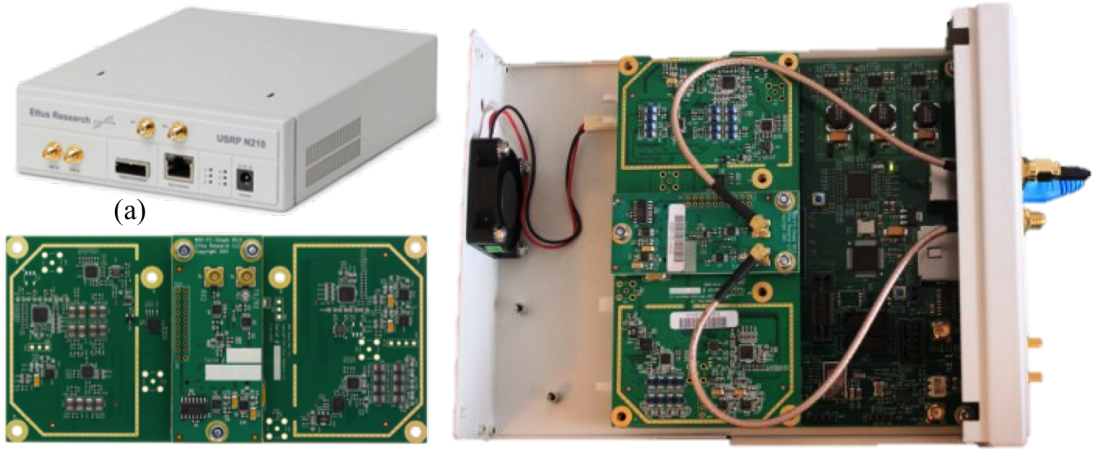


Figure 11.4: Illustration of (a) the USRP N210, (b) the WBX daughterboard, and (c) the USRP related to the block diagram of Figure 11.3. Protruding on the right are the GigE interface (blue), connected to the host computer, and one of the SMA (SubMiniature version A, antenna connector format) connected to DF_1 antenna (gold/black).

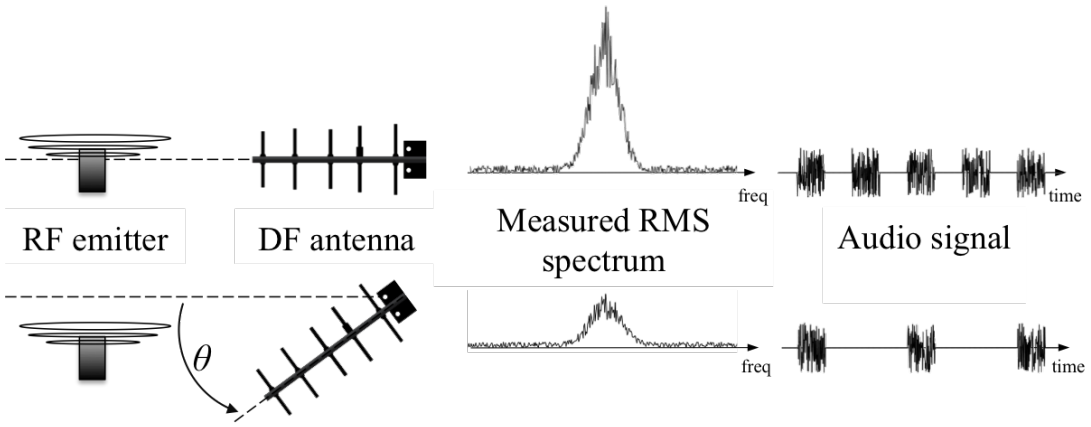


Figure 11.5: Illustration of DF_1 RF Geiger counter implementation. The repetition period of audio bursts gets shorter as the Received Signal Strength (RSS) increases, equating to the directional antenna pointing or progressing towards the RF emitter.

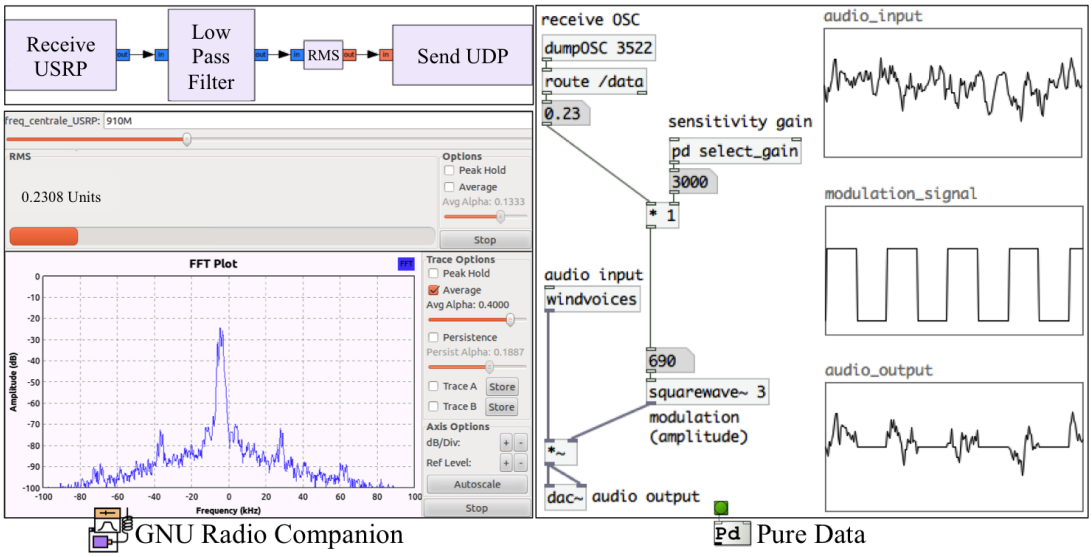


Figure 11.6: Software implementation of DF_1 RF Geiger counter using GNU Radio and Pure Data. The processing chain is first implemented through GRC blocks (top-left), compiled into Python code and executed, which produces the GNU Radio GUI (bottom-left). Measured RMS of the signal received from the USRP is monitored through a level meter (orange bar, RMS at 0.23 "Units") and sent for sonification using the UDP protocol. A Python script (not shown) transforms the UDP messages sent by GNU Radio into OSC messages compatible with Pure Data. Pure Data (right) receives the RMS value to set the oscillation period of a square wave signal (TTL logic, between 0 and 1) used to modulate a noise generator (audio-input).

Table 11.1: Specifications of the components of the PP v1. The DC converter is used to adapt battery's output to USRP power requirements.

Computer	Dell, Latitude E5420, Core i5-2520M (2.5 GHz), 4 Go, 1333 MHz
USRP	Ettus Research, USRP N210 and WBX daughterboard
Directional Antenna	Directional antenna GSM900 8.2 dBi (see Figure 11.2)
Battery	Energizer, XP18000, DC 5 V, 2.1 A / DC 12 V, 2A DC 19 V, 3.5 A
DC Converter	Mascot, Linear Converter DC 12 V / DC 6 V

Pure Data

Pure Data is an open source visual programming language designed to manipulate audio data streams. Well suited to sonification, it is similar to the Max/MSP software used in the studies in Chapters 8, 9, and 10 yet compatible with Linux-based operating systems. It is also based on blocks, each block referring to an internal C method.

For the RF Geiger counter implementation illustrated in Figure 11.6, the measured RMS was received through OSC and used to define the oscillation period of a square wave. This wave then modulated a pink noise generator, producing a sequence of bursts, accelerating as the received RMS increased.

11.2.3 Prototype v1

The first version of the PP was implemented in a ruggedized suitcase, based on a standard laptop connected to a single USRP, itself receiving the signal of a directional antenna. Users wore a headset and manually steered the directional antenna, using a mouse wheel to control the DF sensitivity gain (e.g. to reduce sensitivity as they approached the RF emitter). Table 11.1 lists the specifications of this first version. Figure 11.7b shows some photos of the prototype. //assets/partII/3.1 physical-prototype-v1-principle.m4v and //assets/partII/3.2 physical-prototype-v1-search.m4v videos provide some insight on its behavior for the simple Geiger counter sonification of DF_1, further discussed in Section 11.3.

11.2.4 Prototype v2

Used in field tests with firefighters, the PP v2 was a more compact and lighter backpacked version, based on a small ruggedized computer (MINI PC BRIK 2.5" 3I270D-V4G). To enable further sonification paradigm assessments, a second USRP was added to the architecture of the first prototype, along with a second battery and converter

Table 11.2: Specifications of the RF emitter's components. The 27 dBm (500mW) power of the A/V Transmitter was augmented by the 2 dBi gain GSM Antenna. The resulting 29 dBm was in the upper range of standard GSM900 transmission powers, imposed by the network between 5 and 39 dBm.

Antenna	Adactus, ADA-0186QU, GSM Antenna Penta Stub 2 dBi SWIV
RF emitter	Digital Products, 900 MHz A/V Transmitter, 500 mW DPCAV
Battery	Pro-Tronik, Batterie LiPo Black Lithium 3s, 11.1 V / 2200 mAh / 35 C

needed to power it. The PP v2 could simultaneously process the signal received by two antennas, required for DF designs such as DF_3. The overall design was mounted on a metallic frame to improve its robustness and ensure the system's cooling once inserted in the backpack, illustrated in Figure 11.7c. A typical search with the v2 and the "spatialized RF Geiger counter" sonification of DF_3 is presented in the video //assets/partIII/3.4 physical-prototype-v2-search-2Rx.mp4[2].

The binaural rendering required for DF_3 sonification was processed in Pure Data via the "Earplug binaural synthesizer" [240] (Pure Data add-on library). Except for sonification-related implementation, PP v2 reused both GNU Radio and UDP-2-OSC bridge of Figure 11.6, adapted to process the RMS of the signal received by the second directional antenna.

11.2.5 RF Emitter

Using a real GSM cellphone for the field testing would have required a complete control of its RF layer, i.e. being able to trigger it into emit mode on an isolated frequency. As detailed in Chapter 2, achieving this level of control is possible yet requires the deployment of a complex RF system, exceeding the technical capabilities for the foreseen evaluations. To simplify the procedure concerning DF designs assessment, a simple RF emitter was constructed, based on a 900 MHz Amateur TV emitter. The emitter's specifications and illustration are reported in Table 11.2 and Figure 11.7a.

11.3 DF Performance Evaluation Based on the Physical Prototype

This section follows Section 10.3 from the previous chapter, gathering the observations issued from preliminary tests with the PP v2. The performance and behaviors of DF_1

[2]The pair of directional antennas is hand-held in the video for illustration's sake.

Figure 11.7: Illustration of the RF emitter (a) and the PP v1 (b) and v2 (c). RF emitter sized 7.5cm×13cm×4cm (W×H×D, in cm) and weighted 300 g. PP v1, handle folded, sized 34cm×54cm×23cm and weighted 16 kg. PP v2 aluminum structure sized 28cm×35cm×17cm for a total backpacked size of 31cm×46cm×23cm and weighting 8 kg. (d) Simplified block diagram of both v1 and v2 prototypes.

and DF_3 PP designs, similar to their VP counterparts introduced in Table 10.2, were compared during search simulations in realistic conditions. The general performance of both PP designs were compared with those of three other search and rescue "systems" usually deployed by search party: (1) trained dog, (2) avalanche transceiver, and (3) *geo-stereophone*. Search simulations, observations and discussions were conducted in collaboration with an ergonomist intern, in charge of assessing systems usability throughout the steps of a typical search operation.

11.3.1 Method and Experimental Design

As a reminder, DF_1 sent the RSS measurements of a single directional antenna to a Geiger counter sonification while DF_3 employed a pair of directional antennas, resulting in a similar RSS sonification coupled with a left/right panning based on the ratio of both antennas RSSs. The two DF designs were implemented with the PP v2 illustrated in Figure 11.7c. The same spearcon-based sonification feedback as in Section 10.3 was provided over the RSS sensitivity gain levels, controlled with the wheel of a bluetooth mouse strapped to user's leg. The video //assets/partIII/3.3 physical-prototype-v2-search-1Rx.mp4 illustrates a typical search sequence with DF_1 in an indoor/outdoor environment. Video //assets/partIII/3.4 physical-prototype-v2-search-2Rx.mp4 illustrates a search sequence with DF_3 in an outdoor environment.

Four subjects participated in the experiment, amongst which two firefighters from a SDIS search and rescue squad. For each DF design, subjects had to search for 2 trapped victims, equipped with the RF emitter of Figure 11.7a, in the 100×50 m^2 pile of rubble illustrated Figure 11.8. After a brief training where they were introduced with the PP controls and sonification paradigms, subjects started their first search. The search was achieved when subjects visually located the trapped victim or when they were definite on its location when the trapped victim could not be seen without digging (completely buried in the pile). Subjects were then asked to await, away from the rubble pile without listening to DF output, while a new victim took its position in the search environment.

The results discussed below are not issued from formal measurements, but rather from subjective assessment of participants performance by external examinators. These search tests were intended as a pre-evaluation of the different audio DF designs performance. As a future study, this experimental design could be extended (more subjects, quantitative logs, etc.) to allow rigorous comparison of the considered audio DF designs.

11.3.2 Results and Designs Comparison

Searches with DF_1 proved rather efficient. The four subjects took only a few minutes to master the Geiger counter sonification and its gain sensitivity level selection. Trapped

Figure 11.8: Illustrations of the tests sessions with the SDIS. Comparison of search performance between DF_1, geo-stereophone and dogs. Victims were trapped under the rubbles, equipped with the RF transmitter pictured Figure 11.7a.

victims were localized in $\approx$ 4 min in average, performance comparable to that of the visual DF presented in [19]. One of the test subjects did not instinctively steered the directional antenna towards the ground in the last stage of the search, i.e. at less than 5 m from the target. This resulted in several back and forth above the actual victim's position as the subject complained that the DF "randomly" switched from high to low RSS while pointing in a constant direction.

DF_3 slightly outperformed DF_1 in terms of task execution time. Users judged that search with DF_3 allowed to focus on the navigation rather than on RSS maximum location as it was enough to "make sure that the sound remained in-between the ears" (i.e. same RSS for both $\pm 45^\circ$ directional antennas) while progressing in the search environment. As the sound progressively shifted to the left or to the right, users naturally compensated by slightly turning right or left without pause in their progression. Compared to DF_1, this feature proved helpful in situations where untrained subjects, after a first decision on target's direction would simply walk straightforward, without further adjusting their estimation (e.g. without briefly scanning left and right to make sure they were still aiming at the maximum RSS). The subject previously confused with DF_1 did not reproduce the same mistake with DF_3. As said subject experimented DF_1 condition first, it was not clear whether this was related to basic learning effect or to DF_3 paradigm. It is believed however that the dual antenna combined with sound spatialization could help DF users' interpretation of such situations, e.g. when they just walked by the victim's location as DF_3 produced a characteristic shift in sound position followed by erratic left/right oscillations (see end of //assets/partIII/3.4 physical-prototype-v2-search-2Rx.mp4). The assets of spatial diversity in sonification related to antennas positioning are further discussed in Chapter 13.

As observed during the VP assessment presented in Chapter 10, both DF_1 and DF_3 designs were subject to the uneven dynamic of the inverse square law on RSS as a function of DF-to-target distance. As for the VP implementation, providing a manual control over the sensitivity gain allowed subjects to partially overcome this issue after a few minutes of training. Discussion with users and ergonomist led to believe that the manual Geiger counter sonification scaling and its associated spearcon-based feedback as implemented in both videos //assets/partIII/3.4 physical-prototype-v2-search-*Rx.mp4 should be improved to improve distance estimation for untrained users. The four subjects agreed on being able to estimate when the target was close by (in a $\approx$ 10 m radius) yet were not confident regarding absolute distance estimation during most of their navigation.

During calibration tests preceding this experiment, it appeared that one of the most time-consuming stage of the search with the PP was the identification of meaningful variations in RSS leading to a first DOA estimation (for distant targets). This "catching" stage issue is well known amongst avalanche transceiver users [241]. Through these tests, we would often hopelessly walk just a few meters in each direction before giving up the search. With DF_1 implementation, a trained user could catch the signal at $\approx$ 250 m of the RF emitter in free field scenarios, indicating a sensitivity subsequent to the receiving antenna, coupled with the sonification scaling, of ≈ -43 dBm (see Appendix Section 11.5 for detailed calculation). Catching the RF signal was not a critical issue in the 100×50 m^2 search environment considered in the present study, yet often hindered subjects mishandling DF sensitivity gain levels unable to perceive any clear RSS maximum when using too low a sensitivity gain in the initial stage of their search (away from the victim). This maximum catching distance would obviously decrease for search in indoor environments due to the presence of obstacles increasing the path-loss coefficient between RF emitter and DF receiver. Search situations where victims are deeply buried under rubbles or trapped behind blocks of concrete will also impact DF maximum range [242].

The PP sensitivity could be improved, e.g. following the implementation described in [19] where designers used a cascade of low noise amplifiers and a high quality directional antenna (17 dBi gain) to cover larger search areas. Whatever the DF sensitivity range, a procedure similar to the cross-like or perpendicular bisector techniques [243] already employed with avalanche transceivers seems appropriate for the initial stage of the search to help users to focus on a systematic task until the DF provides meaningful information.

Worth noticing, subjects using DF_3 seemed more confident in assessing whether or not the DF received any signal at all from the RF target: in the absence of any significant signal, the perceived sound erratically oscillated between the left and right ear, stabilizing only when the signal became strong enough to stabilize the RSS ratio

between the two antennas. During calibration tests with DF_3, it was also observed that sound position as a sonification parameter could easily be ignored to focus on tempo, pitch, etc. This feature could prove useful in indoor environments where left/right shifts may no longer be significant because of dense multi-path propagation conditions.

11.3.3 Observations on Design Ergonomics

Through successive tests, It progressively appeared that users confidence in the DF behavior was a critical issue. As said confidence can be undermined in a matter of seconds, particularly in stressful search conditions, one must anticipate potential situations where users would no longer feel they can rely on the DF for navigation. Both technical improvements and predefined procedures can be employed to make the DF more robust to these situations, as the cross-like technique evoked to assist users during the "signal catching" phase.

While positioning the DF antenna on users' head allowed them to keep their hands free during the navigation, subjects agreed that a hand-held solution would be more convenient for precise manipulation of the antenna during the last stage of the search. Discussed in Chapter 13, an in-between solution would be to design a DF based on an antenna strapped to firefighters' helmet that can be removed at will when users need to scan a specific zone of the search area.

Finally, it appeared during the calibration tests that search situations involving both indoor and outdoor environments where the most complex to deal with. Users had to sometimes make arbitrary decisions on a target's potential location: whether it was inside the building or behind it, in the next room or outside, etc. The search scenario of //assets/partIII/3.3 physical-prototype-v2-search-1Rx.mp4 (search from indoor to outdoor, target near to one of the building's walls) often resulted in a lengthy search inside the building before any attempt was made to locate the target from the outside. Indoor situations also raised some issues regarding up/down estimations, where users spent non-negligible amounts of time searching for the RF beacon in the room just bellow/above its actual location. Before training, users did not instinctively aim the DF at ceilings or floors but rather kept steering the directional antenna in the horizontal plane. The left/right spatial diversity employed in DF_3 could be extended to compensate for up/down uncertainties, as proposed in the DF design presented in Chapter 13.

11.3.4 Comparison with Existing Search Solutions

Simultaneously with the search simulations conducted with both DF_1 and DF_3, three other systems were used to search for trapped victims using (1) trained dogs, (2) avalanche transceivers, and (3) geo-stereophone.

The avalanche transceiver, a Mammut Arva Barryvox Pulse model[3] was used with its audio interface only, consisting of a Geiger counter sonification with discrete states (e.g. sudden shift from 60 bpm to 120 bpm as the RSS came above a certain threshold). Both DF_1 and Barryvox transceiver were roughly equivalent in terms of task execution time, while the small packaging of the transceiver allowed users to move more freely during the search. For near-field search situations when subjects closed on the target, the avalanche transceiver outperformed DF_1, due to its secondary search mode automatically triggered a few meters away from the target, i.e. from its tween transceiver. This shift in operational mode clearly informed users that the target was nearby, which instantly modifyed their search attitude. An equivalent mode could be implemented for the PP, assuming knowledge of RF emitter's emission power (controlled by the network for GSM cellphones, see Chapter 2). The diversity of obstacles presents in urban environments and their various impact on RF path-loss would however prevent any reliable distance estimation compared to the uniformness of mountain-like environments where avalanche transceiver are typically employed. Where snow has a known and predictable impact on said path-loss, it is harder to anticipate the nature of obstacles in-between the DF and its target in urban environments, which prevents any reliable RSS-based distance estimation.

Trained dogs clearly outperformed both DF designs, localizing trapped victims approximately twice faster. This result could be partly attributed to the ease with which trained dogs moved through the search environment compared to firefighters equipped with the PP, cumbersome despite its backpacked design. Compared to dogs, their was no time constraint on the DF usage though, and DF search efficiency did not suffer from environmental conditions. Discussions with dog handlers revealed that rescuers cannot expect trained dogs to search for more than 20 min in a raw, as dogs will progressively tire of what they perceive as a "hide-and-seek game". They also reported that environmental conditions such as rain, drifting clouds of gas, or other dogs in the search area could hinder trained dogs performance, or let them completely blind to trapped victims' scent. Trained dogs on the other hand did not require the presence of a cellphone on the trapped victim.

Compared to search with the geo-stereophone, subjects using DF_1 performed approximately 10 times faster, and did not suffered from acoustical noise in the search area (e.g. other squads digging to rescue a victim). The geo-stereophone is a device composed of two microphones, applied on the ground to pickup sounds coming from potential trapped victims. Acting like ears, the inter-microphone level difference can be interpreted by a trained listener to iteratively estimate the victim's location as the microphones are moved through the search area. Its use supposes a team of three firefighters: one listener and two in charge of carrying the microphones, positioned at 5 m on both sides of the listener for spatial diversity. A fourth individual is generally

[3] http://www.mammut.ch.

needed to produce the "call" in-between the searches. Using a metallic pole to hit the ground, the call signals potential trapped victims that they are being searched, and should acknowledge their presence by likewise producing a distinctive sound. While the use of a geo-stereophone does not require victims to carry a cellphone, it works only for conscious victims able to answer rescuers call and will mobilize up to four firefighters, requiring the interruption of any digging activity in the search area.

11.4 Conclusion

This chapter detailed the implementation of a Physical Prototype (PP) of a DF for search and rescue operation. The PP was based on a typical Software Defined Radio (SDR) architecture, which allowed one to handle most of the signal processing and all the sonification routines through software implementations on a computer. A crude RF emitter was also engineered, simulating trapped victim's cellphone to readily test various search scenarios and DF designs without relying on the heavy network deployment process described in Chapter 2.

A preliminary study was conducted to assess the performance of the two DF_1 and DF_3 designs introduced as Virtual Prototypes (VPs) in Chapter 10. The study was designed in collaboration with professional firefighters specialized in search and rescue operations and an ergonomist intern in charge of assessing DFs' usability throughout the steps of a typical search operation. The study consisted in search simulations conducted in a realistic post-earthquake environment, based on both DF designs and three other search systems typical deployed by during search and rescue operations: trained dogs, avalanche transceivers and geo-stereophones.

DF_1 and DF_3 proved rather efficient during these simulations. DF_3 design was judged to offer a slight advantage over DF_1-based searches, due to the immediate feedback it provided when subjects went astray from the correct navigation route, defined by the Direction Of Arrival of the maximum Received Signal Strength. While both DF designs were clearly outperformed by trained dogs in terms of task execution time, they equalled avalanche transceiver's performance and achieved victims' location ten times faster than geo-stereophone. This preliminary study along with calibration tests conducted in indoor/outdoor/forest environments also allowed to gauge the ecological validity of the VP implementation, and to identify situations where the considered DF designs would fail to reliably assist users in their navigation.

Numerous discussions with the search and rescue squads followed these tests, focusing on their requirements regarding DF performance and range of application. The general ergonomic of their own equipment was carefully studied to understand their needs in terms of size, weight, simplicity, packaging, etc. The results of these discussions and observations are reported in Chapter 13, alongside an overview of the

principles and guidelines that emerged from the research on audio DF design discussed throughout this manuscript.

Future work would involve the implementation of a PP for automatic audio DFs, to enable the performance assessment of DFs based on DOA estimation, such as the DF_2 design introduced in Chapter 10, in realistic search conditions. For a SDR-based architecture, this implementation would require the conception of a RF switch [237, 238] or the synchronization of several USRPs [244, 245] to coherently process the signals received by the M antennas of the DF array. Another solution would be to design a computer-less architecture, i.e. based only on FPGAs (Field-Programmable Gate Arrays) and DSPs (Digital Signal Processors) components, reducing the PP implementation cost and size at the expense of its malleability as the signal processing and sonification routines would have to be developed through low-level programming languages. Finally, the discussion of Chapter 6 suggests that the implementation of an automatic DF compliant with portable applications would require a parallel study to improve state-of-the-art arrays compactness without reducing DOA estimation robustness to multi-path propagation.

Based on this improved PP, an exhaustive study on DF_1, DF_2, and DF_3 performance in realistic search conditions would provide some insight on whether or not automatic DF designs significantly outperform manual's, and whether or not DF_3 sonification and spatial diversity outperform DF_2 operational readiness and implementation complexity.

11.5 Appendix: Calculation of the PP Sensitivity

With DF_1 implementation, a trained user could catch RF emitter's signal at ≈ 250 m in free field scenarios using the Geiger counter sonification with the highest sensitivity level. For a distance of 250 m, The free field path loss applied on an electromagnetic wave, noted $PL_{freeField}$, is defined as:

$$PL_{freeField} = 20 \log_{10} \left(\frac{4\pi df}{c} \right)$$

with $c = 3.10^8$ m.s^{-1} the speed of light in air. For $d = 250$ m and $f = 900$ MHz, $PL_{freeField} \approx 80$ dB. Defining G_{Tx} and G_{Rx}, the emitter and receiver gain respectively and P_{Tx} the emitting power, the link budget from RF emitter to DF antenna is defined as:

$$P_{Tx} + G_{Tx} - PL_{freeField} + G_{Rx}$$

which with the 2 dBi emitter antenna, the 8.2 dBi DF antenna, and the 27 dBm 900 MHz Amateur TV emitter power gives a -43 dB sensitivity for DF_1, subsequent

to the receiving antenna (coupled with the sonification scaling), as proposed in Section 11.3.2.

Chapter 12

Virtual and Physical Prototypes Comparison: Ecological Validity Assessment

Contents

This chapter examines the ecological validity of the Virtual Prototype (VP) presented in Chapter 10 regarding the Physical Prototype (PP) detailed in Chapter 11. Said examination supposes the comparison of two different DF designs, both implemented as VPs and PPs. Observing the inter-design evolution in both implementations is thought to provide some insight on how far the results issued from the VP can be applied to improve the PP.

At present, only the virtual part of the evaluation was conducted as a formal inter-participant comparative study. While the suggested methodology demands a companion study on the PP, significant results were still acquired regarding the impact of the Virtual Environment (VE) on VP validity. Part of this chapter has been published in the proceedings of the 20$^{\text{th}}$ International Conference on Auditory Display (2014) [2].

12.1 Introduction

The principle concern in the use of virtual prototyping is its realism, or rather the apprehension of its distance from the PP behavior [246–248]. Once this apprehension,

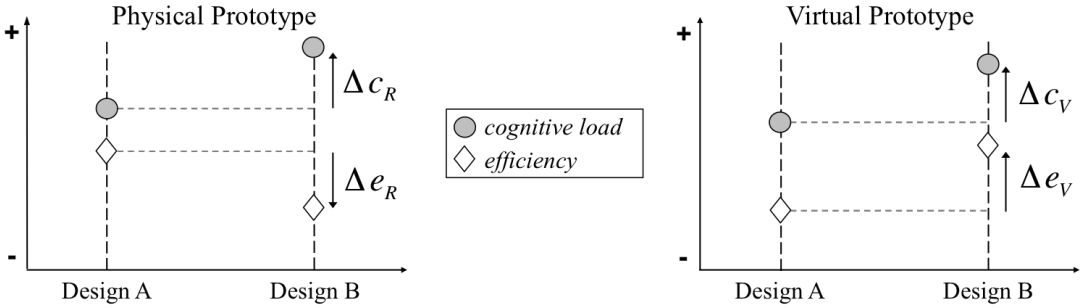


Figure 12.1: Illustration of the suggested ecological validity characterization method. Said method implies to implement two different designs (A and B) as virtual and physical prototypes. On the figure, Δe and Δc respectively represent the increase in efficiency and cognitive load between design A and B. Along with a rigorous qualitative assessment, the correlation between Δc_R and Δc_V or Δe_R and Δe_V informs on the ecological validity of the implementation. In this illustration, the core VP implementation appears reliable in terms of cognitive load while it does not reflect the efficiency decrease observed between PPs A and B.

namely the characterization of its "ecological validity" is achieved, the VP can be used to improve only the design aspects it reliably reflects. Naturally, part of this characterization will concern the impact of the VE on task execution and user behavior [249]. A known method to assess the validity of a specific VP implementation is to compare it to an equivalent PP in terms of performance and ergonomics [250]. As the prototype is modified in both environments (real and virtual), one may observe the correlation of performance evolution, as illustrated in Figure 12.1.

The designs considered in this study were based on the DF_1 introduced as a VP in Chapter 10 and a PP in Chapter 11. Both designs consisted of a directional antenna whose Received Signal Strength (RSS) was sent to a Geiger counter sonification metaphor, further detailed in Section 12.3.1. To reproduce the method presented in Figure 12.1, the two designs differed only in the placement of the directional antenna:

> Design A: directional antenna **in** the DF user's **hand**,
> Design B: directional antenna **on** the DF user's **head**.

While far from state-of-the-art DF implementations (see Chapter 3), the proposed DF core design offered a simple case study of low implementation and human-machine interaction complexity. Coupled with the hand/head minimalistic inter-design modification, the experimental design focused on limiting the amount of independent factors involved in the assessment rather than comparing two promising implementations. Initially, understanding the distance between the VE and the reality is a priority over any ergonomic, performance, or sound aesthetic considerations.

Table 12.1: Independent and dependent variables of the experimental protocol.

Independent variables		
Participant	10	random variable
DF design	2 (conditions)	DF_{Hand}, DF_{Head}
Target	6 (iterations)	T_1, T_2, ... , T_6

Dependent variables
task execution time, covered distance, average speed
move box* orientation (relative to VE or target)
DF antenna orientation (relative to VE or target)

* represents user's viewpoint within the VE

12.2 Experimental Design

The experiment consisted of a DF assisted target localization task carried out in a CAVE (Cave Augmented VE). It involved virtual targets disseminated in a VE which participants had to gradually progress towards and localize using one of the two considered DF designs. A total of 13 volunteers (aged 25 to 40 years old) participated in the experiment. Three out of the 13 volunteers did not complete the experiment, subject to cybersickness induced by the VE [251]. Participants were introduced to the task after answering a set of questions on their experience relative to the considered experiment (VE practice, use of a DF, etc.). They had a varied range of previous experience in sonification and Virtual Reality (VR), from none to expert, while none had ever used any DF-like apparatus. All participants reported normal vision (or corrected to normal), hearing, and physical condition.

During the experiment, participants performed two sessions, one with each design, involving the search of six targets each (repeated-measures design with DF and target conditions as fixed factors). DF design conditions and target orders were evenly balanced across subjects to avoid potential carryover effects. The only instruction was to perform the search as fast as possible. Communication between the participant and the experimenter was not allowed during the search. Instead, questions and comments were heartily encouraged during the training session and the post-experiment interview. The Witmer and Singer presence questionnaire [252], along with open questions during the pre/post interviews were used to detect VE related issues (cybersickness, VE malfunctions, confusions, etc.). The experiment typically lasted one hour, except when participants asked for additional pauses between sessions. Table 12.1 presents the evaluation metrics employed for quantitative analysis of the experiment's results, concerned with participants behavior and performance during the search.

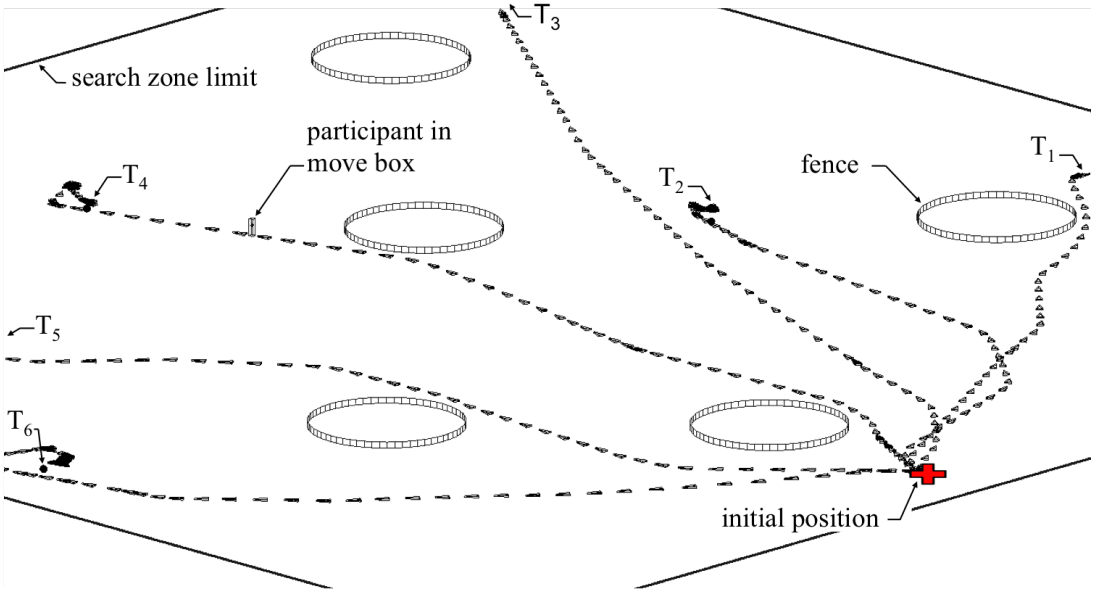


Figure 12.2: VE and target localization task (one session of six sequential targets) illustration. Dotted lines (succession of triangles, see Figure 12.7) represent the move box (CAVE point of view / position in the VE) paths in the VE.

12.3 Experimental Setup

The experiment was conducted in the EVE architecture described in Section 10.2.3. The virtual scene was created in Blender, deployed in EVE with BlenderVR, its adaptation to virtual reality architectures [3] (see Appendix A). The sonification was handled in Max/MSP, synchronized with Blender through OSC messages. Audio feedback was presented via a Radio Frequency (RF) wireless headphone module (Sennheiser EK 2000 IEM and HD570 headset). To provide some ecological equivalent of walking, navigation in the VE was based on the hybrid Walk-In-Place metaphor detailed in Section 12.3.1. DF related interactions employed a 6 DoF tracker that served as a virtual antenna for the hand steered design. The virtual head-mounted antenna was associated to the tracking glasses participants wore for 3D stereoscopic rendering. A 8 kg backpack was worn by every participant to represent PP related load.

12.3.1 Experimental and Audio-Visual Stimuli

Virtual Scene

The virtual scene was a close reproduction of the 6 ha square field identified as suitable for a companion PP experiment. The field was populated with circular fences of variable size to simulate inaccessible areas of the real environment (see Figures 12.2), obstructing participants' path and forcing them to modify their search patterns. An

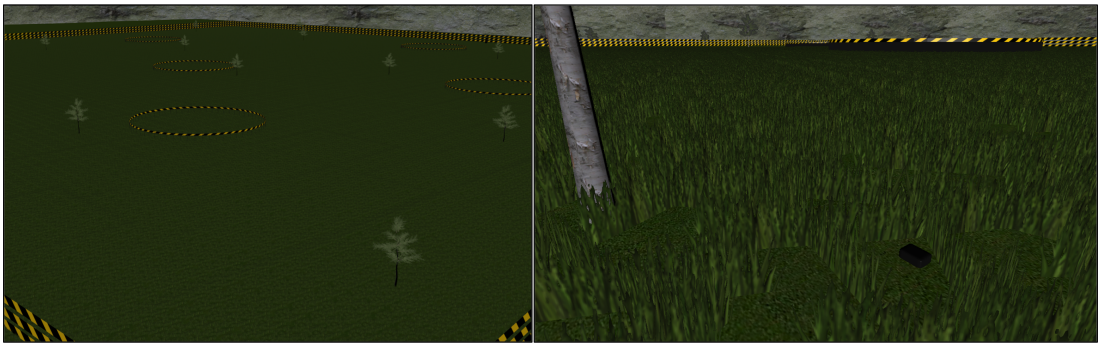


Figure 12.3: VE screen shots. (left) Aerial view and (right) user point of view (front screen only). The black box on the right represents a RF target.

impassible border barrier established the search area limits, illustrated on the left of Figure 12.3. Targets were represented by wallet sized boxes ($14\times10\times2\,\mathrm{cm}^3$) placed on the virtual scene ground, within the accessible area (Figure 12.3, right). To limit visual clues impacting on the audio aided search task, 20 cm high grass patches were homogeneously distributed in the scene to avoid visual target localization from afar. Visual references were added to the virtual scene (trees, mountain range, etc.) to enhance motion perception and path integration [253].

RF Propagation

RF propagation was first simulated using the IlmProp 2D raytracing model [71] employed in the VP implementation Chapter 10. Due to the simplicity of the search environment (flat field without any RF obstacles), the propagation model was reduced to a simple free space, inverse-square propagation law with no noticeable difference on the VP behavior. The virtual antenna was implemented to match the directional characteristics of the PP antenna detailed in Figure 11.2 Chapter 11. Its output power was continually fed into the sonification algorithm, from BlenderVR to Max/MSP using the Open Sound Control (OSC) protocol. As participants approached the target, the antenna output power increased, resulting in a corresponding increase in the Geiger counter repetition frequency. Participants could then select the DF sensitivity level that presented the most meaningful range of repetition frequencies for rhythm appraisal using the analog buttons on the 6 DoF tracker. For the $\mathrm{DF}_{\mathrm{Hand}}$ design, the 6 DoF tracker served as virtual antenna. For the $\mathrm{DF}_{\mathrm{Head}}$ design, the virtual antenna was simulated based on participants' head orientation.

Navigation

The Walk-In-Place (WIP) metaphor was a modified version of the one introduced in Section 10.2.3. The WIP was based on a Nintendo Wii Balance Board (N-WBB) and

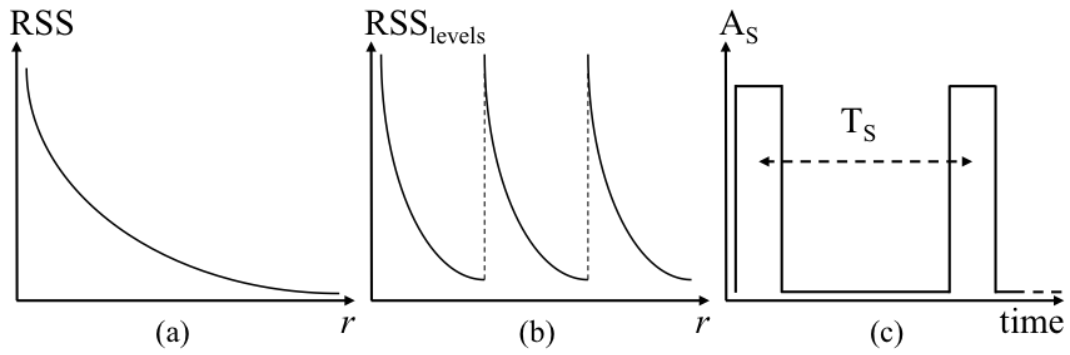


Figure 12.4: (a) Inverse square law illustrating the DF RSS dynamic as a function of DF-to-target distance r. RSS=$G_{(\theta,\varphi)}/r^2$, where $G_{(\theta,\varphi)}$ is the directional antenna gain for DF-to-target angle (θ, φ). (b) Piecewise "linearization" of (a) through sensitivity levels. $RSS_{levels}=G_i \times G_{(\theta,\varphi)}/r^2$, where G_i represents the gain related to one of the 6 sensitivity levels. (c) Amplitude modulation of the Geiger sonification metaphor. Perceived audio bursts are repeated every T_s, where T_s decreases as RSS_{levels} increases.

a Nintendo-Nunchuk (N-N), hand-held joystick that controlled Z axis rotations of the user move box (i.e. viewport) within the virtual scene, as the EVE system is not a full 360° system. The video in //assets/partIII/1.1 virtual-prototype-cave-illustration.mov provides an illustration of the WIP behavior used in the EVE architecture. To compensate for poor N-WBB based WIP precision for small or lateral movements, a second less realistic displacement paradigm referred to as the *hover* displacement mode was implemented. Freely triggered by participants by pressing an analog button on the N-N, this mode allowed slow and precise control of move box translations in the virtual scene based on the N-N joystick[1]. N-N and WBB streamed data to BlenderVR via a VRPN (Virtual Reality Peripheral Network) server through a bluetooth interface. Participants were required to remain on the N-WBB (i.e. not to walk in the CAVE) during the search task.

Sonification Design

The sonification metaphor applied a mapping of the DF RSS to the tempo of a repeated sound sequence. Due to the free field RF propagation model implemented, the RSS followed the dynamic illustrated in Figure 12.4a as a function of distance between the DF user and the RF emitter. To provide the DF user with a roughly linear dynamic, i.e. to perceive a more constant rhythmic evolution as one progressed towards the RF emitter, the function of Figure 12.4a was divided into the piecewise relatively linear function of Figure 12.4b based on the application of decreasing sensitivity gains.

[1]While in hover mode, participants could no longer control the Z axis orientation of the move box.

Switching between sensitivity levels allowed the participants to select a dynamic range well suited to the current DF-to-target distance. This shift was achieved using two analog buttons of the 6 DoF tracker (plus and minus one sensitivity level, see Figure 10.6 Chapter 10). The size of the environment, i.e. the maximum distance between DF and RF emitter, suggested the creation of six different sensitivity levels: a tradeoff between DF usability (ensure there was a *relevant* sensitivity level, regarding rhythmic variations, for every potential DF-to-target distance) and complexity (not too many levels to deal with). To give the user some feedback on the distance to the RF emitter, i.e. to be able to differentiate between sensitivity levels, the impulse sonification sound, a 440 Hz marimba note, was comb filtered according to the current sensitivity level. As a result, the perceived sonification sounded more and more "metallic" as the sensitivity of the DF decreased.

The upper limit of the Geiger counter metaphor tempo was set at 25 Hz (one sample per 40 ms). Once this maximum reached (for a given sensitivity level RSS) a short *click* notification sound was added to the Geiger sonification every 0.9 s. This sound informed the participants that the current level had reached its limit for the current distance and antenna orientation. This state is latter referred to as *clipped*. Such an upper limit avoided users e.g. attempting to differentiate between 8000 and 8600 bps while using a high sensitivity level near the target.

A typical search trial in the VE is shown in //assets/partIII/4.1 virtual-prototype-cave-illustration2.mov (front screen of the CAVE only).

12.4 Experimental Task

Participants started the experiment with a training session which consisted into finding two targets for each DF design. Participants were divided into two groups, each group starting the experiment with either the hand-held or the head-mounted DF design to mitigate potential learning effect. They then started each of the six target localizations (*iteration*) for the two DF designs (*conditions*) at the same position in the virtual scene, illustrated Figure 12.2. Exploring with the DF and navigating via the hybrid WIP, participants would progress towards each target. The hover displacement mode was freely used for slower and more meticulous search. After each target position validation, the move box was returned to its initial position and the DF sensitivity level was reset to one (most sensitive). Once all six targets were located, participants were free to take a short break before proceeding with the second design condition.

To reduce the impact of the the VE interface design, task completion did not require precise localization of the target. Rather, participants were asked to indicate when they though they were sufficiently close from the target ("within what you perceive as 5 meters in the VE"). Hence, they were explicitly asked to memorize the sonification algorithm behavior (adequate sensitivity level, timbre, and rhythm) near the target

during the training session. This 5 m distance was inspired by the change to near-field operation mode employed by the avalanche transceiver discussed in Section 11.3.4, initially designed to avoid situations where participants spent time searching for a target hidden behind a patch of grass.

12.5 Results and Discussion

This section presents quantitative and qualitative results analysis concerning performance and ergonomic variations between VP designs. Qualitative observations serve to understand the impact of the VP implementation (VE, VP, and task execution) in order to avoid misinterpretation of results.

12.5.1 Quantitative Analysis

In the following discussion, the significance of results has been assessed using the non-parametric Kruskal-Wallis one-way analysis of variance with a p-value threshold of 5%, since differences in group variances prevented the use of the more traditional one-way analysis of variance.

The first part of the quantitative analysis concerned DF performance related results. Regarding differences in localization efficiency, Figure 12.5a shows a comparison of task execution time for each DF design condition. Results indicate no significant difference between the time related efficiency of the two DF designs, to the point where the average time values for the subsets were nearly identical (to the nearest second). Comparison of the results as a function of presentation order (see Figure 12.5b) shows that there is a substantial influence of learning effect, coupled with significant differences in the amount of improvement for each DF design between sessions 1 and 2. The head design appeared initially harder to use (session 1) but showed better results than the hand design in session 2. This could mean that the DF designs differed in terms of potential improvement.

This result is uncertain because of the participant skills repartition with respect to the first design experimented. The term "skill" refers to one's ability to learn and assimilate interactions required by the task at hand (WIP, joystick control, DF usage, etc.). A skilled participant would, amongst other things, present a bellow average total task execution time. Regarding task execution time ranking, the top four participants started session one with the same DF_{Hand} design. Inspection of results in Figure 12.5b, comparing [session 1 DF_{Head} - session 2 DF_{Hand}] and [session 1 DF_{Hand} - session 2 DF_{Head}] indicates a similar influence of the learning effect between sessions for both groups, and the uneven skills repartition that may explain the observed difference in potential improvement.

Thorough investigation of participant questionnaires and oral interviews did not

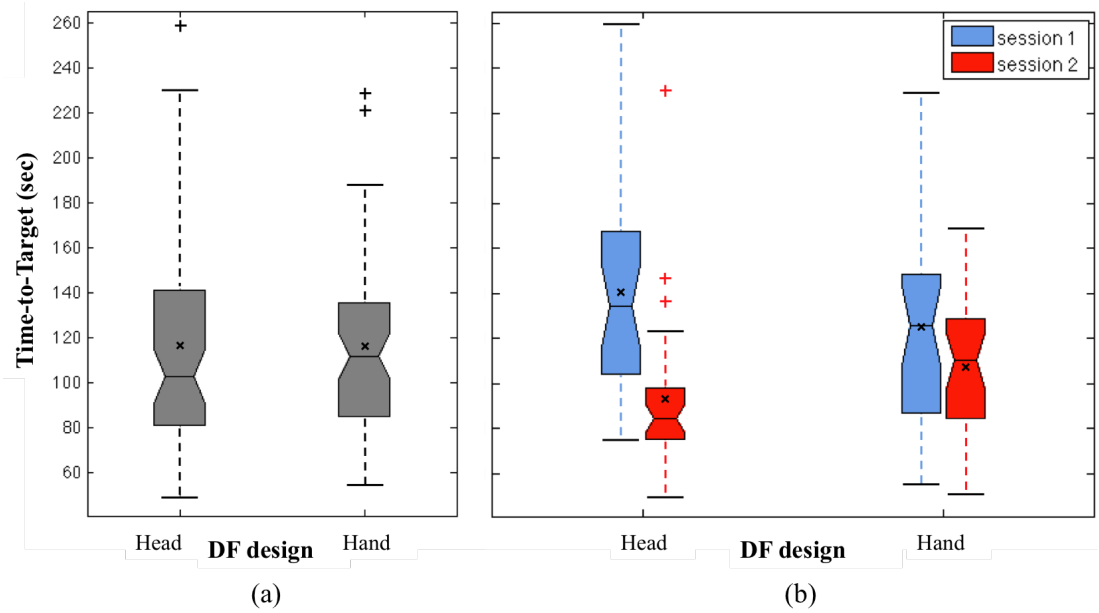


Figure 12.5: (a) Distribution of recorded time-to-target for all participants (10×6 targets) as a function of DF design. (b) Time-to-target results separated by session. "×" and "+" symbols stand as subset means and outliers respectively (with a maximum whisker length of 1.5×inter-quartile range). Notched boxplot middle line indicates the subset median value.

provide any differentiating factor between the highlighted groups. Performances seemed to mainly depend on their ability to handle the VE interface and their involvement in the task (see Section 12.5.2). Participants did not report any significant preference towards either design. At this point, a quantitative analysis of only session two (removing learning effect) would not be statistically significant, leaving only six targets times five participants per DF design.

The inter-target performance analysis indicated that participants took significantly less time to find distant targets than close targets, for target localization times normalized by target distance from the initial position in the virtual scene. There was no significant difference between inter-target covered distances (again normalized) during the search. Participants actually spent in average a sixth of their time exploring within a short distance of the targets (with no significant inter-session, design, nor target fluctuations). This region, defined by a radius of 5 m, is referred to as the "near-field" search area. Observation of the average speed of the move box corroborates the time consuming aspect of the near-field stage of the search.

The ecological validity of this result is difficult to assess without the figures of the companion study with the PP. The calibration tests with DF_1 PP in a free-field

environment discussed in Section 11.3.2 suggested that the near-field search was not particularly time consuming when the target laid visible on the ground. The near-field search also forced participants attention on the unrealistic details of the VE implementation (hybrid WIP displacement metaphor, graphics, etc.). Participants would, for example, report that adding ground bumps on the scene's floor or allowing physical displacement in the CAVE would have considerably helped the near-field search.

Regarding the accuracy of participant estimations and their related efficiency in navigating through the virtual scene, the difference between total covered distance, average DF antenna orientation, and move box orientation (relative to target) were examined for both DF designs. To simplify the analysis, data related to near-field search were not considered, to avoid continuous 180° jumps in orientation as participants walked past a target. There were no significant variation of either distance nor orientation related results between DF conditions. Inter-target, participant, or design localization accuracy analysis, i.e. distance between move box and target when participants validated the target position, did not yield any significant result either.

As a measure of how much participants steered the antenna independent of the move box orientation (i.e. head or hand antenna movements) for each design during the localization task, the metric considered was the sum of angular movements of the DF antenna during the search, defined as

$$\sum_{n=1}^{N_{tot}-N_{near\text{-}field}} abs(\theta_{\text{antenna2moveBox}}[n] - \theta_{\text{antenna2moveBox}}[n-1])$$

$N_{near\text{-}field}$ was subtracted from the total number of recorded samples N_{tot}, removing the near-field phase of the search from the calculation to avoid meaningless 180° shifts. Figure 12.6 reports the statistics of this metric regarding DF designs and sessions. Figure 12.6a indicates a significant difference which supposes that participants scanned their surroundings more thoroughly in the DF_{Hand} condition. Separation of this analysis by session (Figure 12.6b) supports this observation, furthermore suggesting that skilled users (i.e. the [session 1 DF_{Hand} - session 2 DF_{Head}] group, introduced earlier in this section) needed significantly less antenna movements to find the targets in the second session.

Regarding this issue, the VE implementation was believed to induce non-realistic behaviors, and a potential bias on inter-design and comparison across participants. With the head design, participants did not scan (i.e. look) beyond the field of view of the projection screens while rear orientations were scanned with DF_{Hand}. Also, the WIP allowed for a natural dissociation between the walking direction and the head orientation for participants comfortable with the VE interface, while others had to stop walking to scan their surroundings when the directional antenna was positioned on their head (see Figures 12.7b-c).

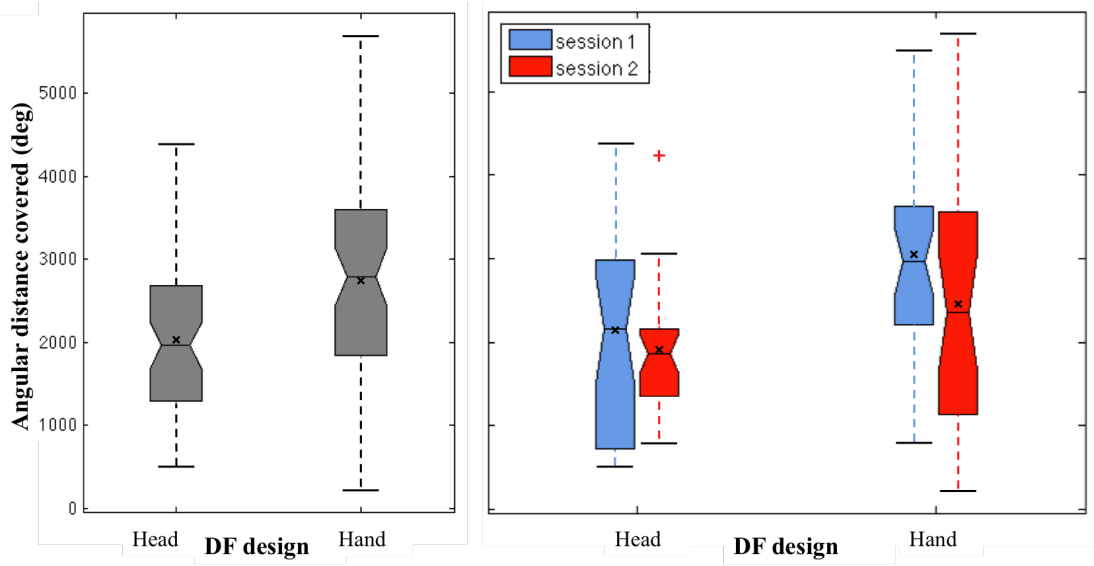


Figure 12.6: (a) Distribution of summed angular distance covered by the DF antenna during the search (10×6 targets) as a function of DF design. (b) Same as (a), separated by session.

12.5.2 Qualitative Analysis

This section is concerned with the observed search strategies, along with non-ecological aspects of the VP implementation.

In the experiment, search strategies appeared to primarily depend on participant skill and their control of the VE interface rather than the specific features of each DF design. The strategies developed in session 1 (for a given DF design) were often reused and refined in session 2 for both groups. The following is a synthesis of observations on participant search strategies during the course of the experiment.

Maximum Power search strategy: involved steering the DF antenna until an angular maximum of signal power was found (i.e. the antenna orientation producing the shortest audio sample repetition period for a given sensitivity level). Every participant started the experiment using this strategy while only half of them exclusively used it until the very end. This strategy was optimal for those able to quickly assess rhythmic fluctuations while advancing through the VE.

Minimum Power search strategy: involved steering the DF antenna until an angular minimum of signal power was found (the reverse of Maximum Power search strategy). This strategy took advantage of the directional antenna's main null (oppo-

site to its main lobe, particularly narrow for the simulated antenna). This strategy was instinctively adopted by three participants, proving to be more time consuming because of the slow dynamic of rhythmic feedback in the low power regions (having to compare between *slow* and *very slow* pulsations). It eventually resulted in more confident estimations for those who were uncomfortable with detection of subtle rhythmic fluctuations. The minimum power search strategy was seldom used in the DF_{Head} condition, as the antenna orientation most often needed to be opposed to the walking direction.

Triangulation search strategy: involved participants moving in the virtual scene without steering the antenna until a rhythmic fluctuation was perceived due to the distance variation between the RF emitter and the DF antenna. The directional characteristics of the antenna were not used: the user moved in the VE to gather information on the target direction. While not optimal in the overall experiment, this strategy proved extremely efficient for near-field searches, i.e. for precise target localization where small movements of the move box produced important shifts in the received signal power (due to the RF propagation model, despite the bin-wise pseudo linearization). For far-field search, participants always coupled this strategy with another search strategy.

Interval search strategy: (*unexpected at first, this strategy was related to the sonification clipping paradigm evoked in Section 12.3.1*) involved steering the DF antenna using a hypersensitive level, i.e. subject to *clipping*. As a reminder, the sound related to the clipping was added to the sonification only for received signal powers above a RSS threshold relative to a given sensitivity level. Considering the antenna directivity, this threshold can be seen as a cone drawn towards the target with its tip at move box position. The clipping sound was heard only when the antenna aim was inside this cone, with an angle depending on the target distance and the current sensitivity level. Identifying right and left clipping limits allowed participants to precisely estimate the target direction. Three participants adopted this strategy, which resulted in slow yet confident estimations. The interval search strategy particularly hindered participants' movements, as the clipping limits (right and left) had to be estimated both at a given distance from the target.

While reasonable performances could be obtained using one of these strategies, the two best performing participants (in terms of task execution time) were able to seamlessly use all four of them during the experiment. As the amount of information increased, they roamed the VE more confidently, adding motion dynamics to the audio feedback (the faster they moved towards an RF target, the greater the Geiger counter rhythmic fluctuations).

Mishandling of sensitivity levels proved to be one of the main performance issues. Participants underestimating the required sensitivity level (for a given target distance)

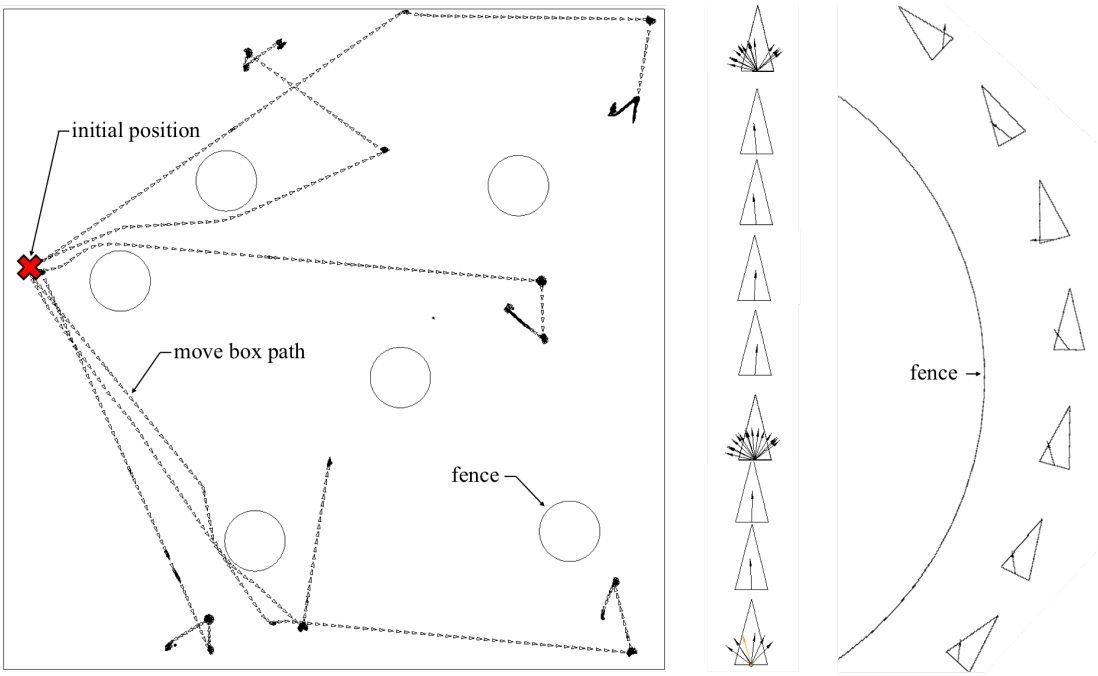


Figure 12.7: (a) Full search path illustration of the first session of one of the slowest participants (regarding task execution time). Triangles represents the recorded move box positions, arrows the antenna orientations. Path straightness and static rotations of the move box (instead of the directional antenna) suggest a non-optimal use of the WIP metaphor (compared to the smooth trajectory curves in Figure 12.2). (b) (zoom of (a)) Temporal dissociated walk and antenna orientation: the user stopped walking every few seconds to scan with the DF. (c) (zoom of Figure 12.2) Participant using both DF and WIP simultaneously.

would, for instance, have to differentiate between a slow and a very slow pulsation (e.g. 0.5 and 0.3 bps) to find their way towards the next target, where an appropriate sensitivity level would have resulted in a comparison between 25 and 1 bps for the same RSS. All three removed participants experienced such a situation, reporting afterward that their irritation and lack of trust in the DF measurements affected their attention in the experimental task. It may be related to their experiencing cybersickness.

The main VE related performance issues came from difficulties using the DF and the WIP metaphor simultaneously. Figure 12.7 illustrates differences in VE exploration efficiency between two typical participants. A suggested improvement would be to thoroughly evaluate participants skills regarding VE related interaction mechanisms to strictly limit the VP analysis to changes related to the DF design.

Ultimately, the task lacked the urgency related to search and rescue operations. This issue was related to the experimental context rather than the VE. While this is an issue regarding ecological validity, it does not necessarily affect the VP external validity [249], i.e. the bias induced thereby does not necessarily invalidate the generalization of the results issued from the VP.

12.6 Conclusion

This chapter presented a study designed to assess the ecological validity of the Virtual Prototype (VP). Said assessment implied the execution of two parallel experimental observations of both physical and virtual prototypes during a target localization task, observing the evolution of the performance of both prototypes as the DF design was modified. The virtual component of this evaluation was carried out for two DF designs as a preliminary evaluation and use of VP concepts. The companion PP assessment is left as a future work for this project.

Both DF designs were implemented and tested in a CAVE implementation of the VP, to enhance participants immersion in the task. Ten participants had to use each DF design to localize several targets in a minimum amount of time. The core DF design was based on a single directional antenna whose RSS was mapped to a Geiger counter sonification metaphor. Implemented DF designs differed only in the position of said antenna: in the hand or on the head of the participants.

Raw performances, e.g. task execution time, were identical for both designs. The individual search strategy adopted by each participants appeared to be selected on the basis of listening skills and participant familiarity with the VR interface rather than on DF design condition.

The VP implementation itself produced a bias between participants according to their abilities to learn and assimilate non-realistic HCI interactions like Walk In Place (WIP) or joystick control. The implementation may also have induced an inter-design bias, the WIP allowing participants to easily move and look in different directions for a large amount of time (while one of the DF design involved head movements to steer the antenna). Using the VP as a test platform to assess the Physical Prototype (PP) would require thorough training sessions to remove any VE-related impact between participants along with a careful examination of eventual benefits from the implementation on specific prototyped designs.

Informal tests with the PP in a flat field environment suggested that the simple RF model (free field propagation) induced DF behaviors that reflected reality. The VP of more complex search environments will however require a more advanced propagation model. All the search strategies deployed with the VP were also naturally used with the PP, except for the “Interval Search”. This strategy has no validity in a real envi-

ronment as it relied on the deterministic nature of the implemented RF propagation model. Noise in the RSS measured by the directional antenna, as discussed in Chapter 9, would for example continuously shift the interval bounds, preventing any reliable direction of arrival estimation. Finally, the "Minimum Power" strategy proved nearly useless in multipath environments as the minimum RSS was not always opposed to the maximum[2].

[2]see indoor tests with the PP in video //assets/partIII/3.3 physical-prototype-v2-search-1Rx.mp4

Chapter 13

Industrial Perspectives for a Cellphone-Based Audio DF

Contents

More of an engineering report than a formal scientific reflection, this chapter presents a short discussion on direction finder (DF) opportunities for search and rescue application. Along with considerations on cellphone-based DF regarding the current scientific and industrial context, a final DF design is proposed based on the knowledge gathered during the research, interviews with end users, surveys, etc. undertaken in the course of this project.

13.1 The Use of Cellphones as RF Beacons

As mentioned in Chapter 2, the network attachment process, required to identify and control GSM cellphones in a given area, requires a governmental authorization. In most countries, said authorization will seldom be delivered to firefighters or emergency response services but for dire crisis situations. A solution that freely allows search parties to use cellphones as rescue beacons hardly seems achievable, more likely to involve some punctual acceptance procedure prior to any deployment. Furthermore, one of the shortcomings of the jamming process described in Chapter 2 is that their is no guarantee that the cellphones forcibly attached on the private network will all belong to a victim, which raises privacy issues regarding unconcerned individuals passing by the search area.

The development of new telecommunication standards is also likely to raise a critical issue regarding activation procedures involving the deployment of a private network in the search area. Compared to GSM, The 3G[1] and 4G[2] protocols for example involve

[1] UMTS, Universal Mobile Telecommunications System.
[2] LTE, Long Term Evolution.

a crossed identification procedure, from network to cellphone and reciprocally. In other words, a 3/4G cellphone will never attach itself to any network not specified on the white list hardcoded in its SIM card. Due to backward compatibility, any 3/4G cellphone currently available on the market will automatically switch back in GSM mode when no 3/4G network is available. While this feature allows to bypass the non recorded network issue, their is no guarantee that this will forever be an available option.

13.2 Opportunities in an Overcrowded Market

Dozens of DF solutions have already been designed that exploit cellphones as localization beacons, not to mention solutions based on cellphones geolocation (favored for distant recon operations). A short reading through scientific papers [19, 81], patents [254, 255], complemented by the analysis of currently available commercial solutions [256–258] clearly indicates that their is not much room for technical innovation regarding portable DF design for search and rescue applications. As the ultimate goal is to locate individuals, any new solution would furthermore have to compete with increasingly efficient location techniques based on organic victims location (smells, heat, sounds, etc.).

In short, there are three key components for every location system: performance (reach, precision, speed, ergonomics, etc.), logistic (deployment, availability, percentage of localizable victims, maintenance) and cost (acquisition, subscription, training, etc.). Naturally, the requirements for each of these components vary with the targeted users. For firefighters and emergency responders, the minimum requirements to commercialize a mobile DF solution would be:

Performance: covers a minimum area of 50 m radius in worst case scenarios, allows to locate victims with an accuracy inf. to 3 m or to define a 3 m radius search area for buried victims. Search time inf. to 5-10 min (from signal detection to final decision, prior to digging), less than 10-20 min training to be able to operate the DF.

Logistic: minimum of 45 min battery operation, ruggedized, operational deployement inf. to 10 min. Minimal maintenance (e.g. functional after more than a month in the response vehicle), allows to localize more than 70% of the victims in the search area. Maximum of 15-20 kg for a backpacked solution.

Cost: varies with design performance. Independent DF modules carried by emergency responders should not exceed 5 k€.

Were all these requirements fulfilled, the commercialization of a new DF would still involve much efforts to find a niche in this already dense market (particularly regarding the limited number of potential customers). Any successful strategy is likely to rely

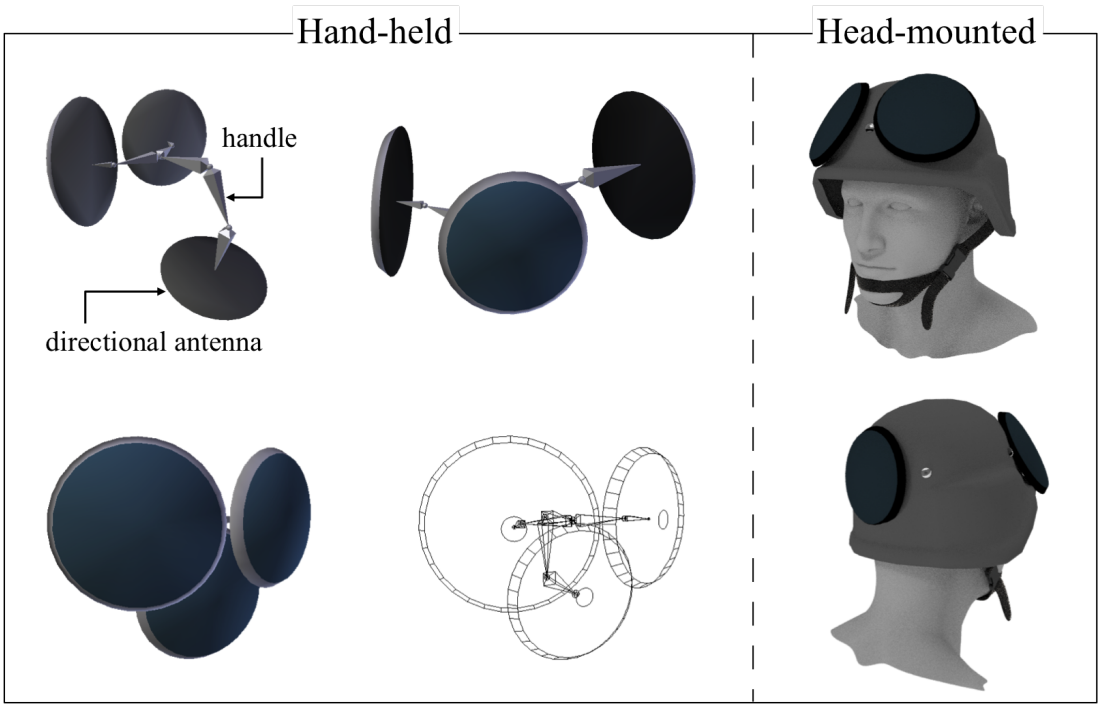


Figure 13.1: Final Audio DF design proposal. The 3 directional antenna are mounted on an articulated structure that can be strapped on firefighters' helmet. Changing the structure configuration, users can adapt the design to several type of search condition. Illustrated antenna is low profile with high directivity, scaled after the LPA7 18V/132 model from Cobham antennas.

on at least one or two features amongst: performance differentiation, originality (buzz effect), low cost solution, acknowledged designer/reseller.

13.3 Audio DF Design Proposal

This section presents the implementation details of a potential audio DF design for search and rescue operation. Inspired from the various assessments reported in this manuscript[3], this design is based on a set of three directional antennas. Presented in Figure 13.1, the general design resembles the DF_3 introduced in Chapters 10 and 11. In addition to the two main directional antennas[4] which Received Signal Strengths (RSSs) are mapped to a binaural Geiger counter sonification, the RSS of a third directional antenna is exploited in the sonification. This antenna is used to notify users when the RSS is stronger in a direction opposed to the one their attention is focused on.

[3] see Chapters 6, 10 and 11.

[4] Scaled illustration based on Cobham LPA7 antenna: http://www.everythingrf.com/products/panel-antennas/cobham/646-105-lpa7-18v-132.

For the "head-mounted" configuration in Figure 13.1 for example, the third antenna is positioned at the rear of the helmet to rapidly warn users when they search in the wrong direction. The sonification suggested for the RSS of this antenna would be unobtrusive, e.g. overlaid to the main audio stream as a discrete pulsating sound or a change in the timbre of the main Geiger counter sonification as said RSS gets stronger than the average RSS of both front antennas.

All three antenna are mounted on an articulated structure which can be strapped to users' helmet or hand-held, e.g. based on a 3-points magnetic fastening. Helmet-mounted operation is advised for fast and rough estimation, e.g. at the beginning of the search. Hand-held operation can be used for proximity or indoor search, to allow one to thoroughly scan a given area or to quickly check difference in RSS as the DF is pushed through an open door. The articulated structure allows to adjust the orientation of the third antenna, to point up or down according to search condition. When searching through the rubble pile of Chapter 11 for example, the third antenna can be used to inform users they just walked over a trapped victim. The orientation of the two main antennas can be slightly modified (from pointing in the same direction to 90° apart) to provide more of less angle width for the main binaural search. For 1D, focused search, a second operating mode would automatically be triggered when the third antenna is folded to point in the same direction than the two main antennas (Figure 13.1, bottom left). The average value of all three RSS could then be monitored via a single Geiger counter sonification. A fourth directional antenna, mounted on a telescopic arm, could be used to probe through the rubbles for precise localization in the final stage of the search. The DF hardware would be based on an new generation of USRP, e.g. the E310 model[5] strapped on a belt. This new design would avoid the use of an external computer as the E310 has its own audio output and comes with an FPGA powerful enough to handle both signal processing and sonification routines.

The DF volume is controlled via a continuous analog switch, its sensitivity level via a discrete one. Changing the sensitivity level triggers a spearcon-based notification and changes the Geiger counter timbre, merging the sonifications proposed in //assets/partIII/3.4 physical-prototype-v2-search-2Rx.mp4 and 4.1 virtual-prototype-cave-illustration2.mov. Two or three different sets of sounds could be selected to fit various environmental noise conditions (from smooth to sharp sounds to avoid environmental sounds masking the sonification). The DF must always produces some audio feedback to comfort users on its operating status, e.g. in the form of a non-intrusive background organic noise to indicate that it is turned on. Audio feedback can be rendered through bonephones[6], to allow users to communicate with other individuals of the search party and keep some awareness of their surrounding environment during its operation.

[5] sized 133×68×26.4 mm for 375 g, http://www.ettus.com/product/details/E310-KIT.

[6] see Section 10.3.

Chapter 14

General Conclusion

Contents

While locating people based on their cellphone is very often seen as intrusive on one's privacy, it however can prove highly useful to assist emergency responders during search and rescue operations. Acting then in a manner resembling avalanche transceivers, cellphones furthermore offer the advantage of being wide spread and daily carried by most people throughout the planet. Numerous designs of Radio Frequency (RF) Direction Finders (DFs), a kind of compass for RF waves, have already been implemented and successfully tested to locate unconscious or lost individuals in crisis situations.

The work presented in this manuscript has been largely influenced by the recent progress of sonification[1] in the design of Human Machine Interface (HMI). While DFs

[1] Audio equivalent of visualization, see Chapter 7.

are usually based on some visual interface to display navigation-related information, the whole purpose of this thesis was to examine the benefits of a sound-based HMI, focusing on sound and perception from the very start of the design process. In an attempt to design an audio-based DF that outperformed visual DFs, the ambition of this thesis was to take advantage of the assets of the human auditory system rather than relying on advanced estimation techniques.

The contributions brought forward by this thesis are presented in Section 14.1. Section 14.2 discusses some of the research perspectives raised by the investigations reported in this document.

14.1 Contributions

The thesis contributions concern the three lines of research respectively discussed in Parts I, II, and III of this manuscript. Section 14.1.1 reports the examination of achievable DF performance in the context of mobile search and rescue operations undertaken in Part I. Sections 14.1.2 and 14.1.3 review the achievements of Parts II and III, related to DF sonification and to the development and use of both virtual and physical DF prototypes for ecological design evaluation.

14.1.1 Investigation on the Achievable Performance of Classical DF Designs in the Context of Search and Rescue Operations

The literature review presented in Chapter 3 confirmed that some Direction Of Arrival (DOA) estimation techniques had already been designed that fitted the requirements of the application context exposed in Chapter 1, i.e. providing robust, fast, and accurate estimations in multipath propagation environments. Most of these techniques however proved to rely on antenna array geometries whose dimensions prevented portable and lightweight designs as discussed in Chapter 5. The various discussions and evaluations that composed the work reported in Part I suggested two potential lines of investigation to optimize DF design for search and rescue operation.

The first option that supposed to address the issue of the minimum DF size by designing a new DOA estimation technique or an innovative antenna array was discarded for the various reasons invoked in Sections 11.2 and 6. A potential bypass was proposed in Chapter 5, supposing the dynamic reconfiguration of a sparse array geometry based on metrics issued from the Cramer-Rao Bound. The considered design compensated for its small number of antennas by a dynamic redistribution of their position to optimize performance regarding DOA estimations on a given region of space. This method proved however inefficient regarding estimation robustness in multipath propagation conditions when coupled with classical DOA estimation techniques.

The second option supposed to start with and improve the performance of a manual

DF design, based on users observing the Received Signal Strength (RSS) of a directional antenna rather than relying on an automatic DOA estimation. The search tests performed with the Virtual Prototype (VP) in Chapter 10 led us to believe that such a simple design was not much less efficient than a more advanced automatic DF. The manual DF designs proposed and assessed in Chapters 10 and 11 suggested that the search efficiency could be increased by simultaneously presenting the information issued from a second directional antenna, taking advantage of the spatial diversity offered by a pair of antennas aimed in different directions. The underlying assumption yet to be confirmed was that the performance of a DOA-based DF could be equalled or even surpassed by a compact design based on two to four directional antennas simultaneously monitored by the user. Rather than providing an estimated DOA, would simple information like "right", "left", "behind", etc. result in an efficient guided navigation.

14.1.2 Sonification Design for an Audio-Based DF Interface

The literature review of Chapter 7 exposed the potential benefits of an audio-based HMI compared to visual displays classically used for DF feedback, not least the natural capacity of the auditory system to locate unseen objects in 360° surroundings. Regarding the assumption on the efficiency of a DF design based on the sonification of multiple antennas' RSS, the literature review partly focused on researches conducted on one's capacity to merge large amounts of information into single sound events and to monitor several auditory streams simultaneously.

The preliminary study presented in Chapter 8 suggested that sonification could handle various kind of DF output, from high level single value DOA estimation to low level mean RSS distributed in space. Repeated listening tests indicated that the perceived prominence of sound parameters should reflect the relative importance of the presented information. The sonification mapping must allow untrained listeners to effortlessly focus on the main source of information but can nevertheless provide a second or third class of information, overlaid on the main stream or in the background of the auditory scene, that users will progressively be able to listen to and take advantage of as they gain in expertise.

Chapter 9 initiated a reflexion on the use of perceptual properties intrinsic to sound parameters in Parameter Mapping Sonification (PMSon). As persistence of vision which transforms static images into animations, the mechanisms of hearing perception, specific to each sound parameter (pitch, tempo, etc.) can be employed to improve sonification designs. In this illustrative study, the natural smoothing effect of inter-onset-interval perception (∼tempo) proved efficient to reduce the impact of measurement noise on the exploration of RSS topologies.

14.1.3 Use of Both Virtual and Physical Prototypes for DF Performance Assessment

The VP detailed in Chapter 10 was employed to assess the behavior and performance of various DF designs in different search environments. The VP implementation was based on the application of a raytracing propagation model to the geometry of a virtual scene, used as the search environment. Said implementation was adapted to a CAVE architecture for lifelike search simulations, leading to the development of BlenderVR, and open-source framework for the development of immersive virtual environments, based on the Blender software. In Chapter 11, several versions of a Physical Prototype (PP) were designed based on the results and observations issued from tests with the VP. The PP was built using a flexible and scalable architecture based on software defined radio, allowing it to handle most of the signal processing and sonification related routines at a software level. Two different audio DF designs, implemented as PP, were assessed and compared to state-of-the-art search solutions during a simulated rescue operation in a post-earthquake environment. Members of a professional search and rescue unit participated in the simulated operation along with an ergonomist in charge of evaluating the usability of the DF designs through the different stages of the search activity.

Chapter 12 presented the first results of an assessment of the VP ecological validity, based on an observation of VP-to-PP correlation over DF design modification. While the full study was not completed, the VP appeared to produce relatively coherent results in most search situations, in light of the comparative evaluations on VP and PP-based DF designs presented in Chapters 10 and 11. Simulations of complex environments such as indoor or indoor-to-outdoor search were limited by the oversimplified propagation model used for the VP implementation. Aside from behavioral evaluations, the VP proved particularly useful for calibrating the various sonification mappings used for the PP as it perfectly reproduced RSS and directional antenna dynamics in free-space propagation scenarios.

14.2 Further Work

The research prospects opened by this thesis are gathered in this section.

14.2.1 Potential Improvements of the Virtual Prototype

Currently, the RF propagation model of the VP was based on IlmProp, an Open Source 2D raytracing model [71]. Aside from being license-free, IlmProp was selected because it allowed the use of the virtual scene geometry as an input for the propagation simulation. A more advanced 3D model such as iBuildNet [259] or the beam tracing

based implementation presented in [260] would be required to refine the ecological validity of the VP and further test DF behaviors in indoor/outdoor environments.

To exhaustively reproduce the considered application context, the impact of other RF emitters on DF estimations, potentially interfering with the targeted cellphone would have to be assessed. Integrating the overall network attachment procedure exposed in Chapter 2 and simulating concurrent GSM networks present in the search area would allow one to validate and improve the logistics of the complete cellphone-based localization process.

14.2.2 DF Distance Estimation and Distance Sonification

The benefits of distance estimation were often evoked during DF field tests with response teams, particularly regarding victims potentially buried under boulders, rubble, concrete blocks, etc. Compared to classical avalanche transceiver where the only material to consider is snow[2], distance estimation with a DF would require some knowledge of obstacle composition, as said composition will define the pathloss imposed on RF propagation. As most firefighters in these response teams are expected to recognize various materials at a glance for safety-related issues, a semi-automated distance estimation procedure might be considered. Supposing the knowledge of cellphone emission power (defined by the host network), a simple interface could be designed to adjust the DF sensitivity to the observed obstacles prior to cellphone distance/depth estimation, based on pathloss information related to the identified materials [242]. Their is still some research to carry out regarding distance sonification paradigms, as the performance of the mappings currently proposed in the literature (e.g. [261]) have yet to match the precision and intuitiveness of even the most basic seven-segment display.

14.2.3 Attention Triggered Auditory Zoom

This concept was first evoked during our designing of the sensitivity levels to compensate for the inverse square-law dynamic of the directional antenna RSS as a function of DF-to-cellphone distance for the ecological assessment experiment presented in Chapter 12. As a reminder, these sensitivity levels were required as the variations of the sound parameter used for RSS sonification were far from being linear as the DF approached the cellphone, not perceivable from afar or far too large near the cellphone.

An obvious solution to avoid the definition of any sensitivity level in the visual domain is to use a numerical display with several seven-segment counters. As their attention naturally focuses on the one that varies "at a reasonable speed", users can easily appraise the variation and its scale, similar to watching seconds on a chronometer rather than minutes or milliseconds. The closest example found in the literature

[2]Distance estimation is indeed available on most avalanche transceiver models.

would be the "Multi-Band Frequency Modulation" proposed in [195], c.f. audio track in //assets/appendix/Multi-Band Frequency Modulation.wav. Applying a frequency modulation with various gains on multiple tones, the authors obtained several levels of variation speeds that resemble this chronometer paradigm. A preliminary study was conducted with three Shepard tones [262], representing the variation of units, tens and hundreds of a counter. As the counter speed increased, listener's attention seemed to naturally shift from one tone to the next[3] to monitor the counter value with the most relevant auditory zoom. A complete experiment would be required to assess the maximum number of tones that can be simultaneously monitored[4] and the efficiency of such a sonification for absolute distance estimation.

14.2.4 Definition of a JND-based Scaling Methodology for PMSon

PMSon designs are often based on arbitrary sonification scaling. This does not represent any issue for most studies, yet can prove problematic when comparing the performance of two or more different sound parameters mappings to convey a given information. As for the PMSons assessed in Chapter 9, a solution would be to clearly establish a scaling methodology based on Just Noticeable Differences (JND) for each sound parameters. Concerned with pitch, inter-onset-interval, loudness, brightness, etc. and potentially gathering results published in early psychoacoustic studies [93, 201], such methodology would allow one to systematically establish a one-to-one JND scale between sound parameter variations to support inter-PMSon comparisons and results integrity.

14.2.5 Formal Evaluation of the Audio DF PP Performance

To truly achieve this work, the performance of the design proposed in Chapter 13 would have to be formally compared to concurrent methods during real search and rescue operations, such as visual-based DFs designs (based on DOA estimation [81] and RSS measurement [19]), trained dogs, geo-stereophones, avalanche teansceivers, etc. Said comparison should focus at least on raw performance, robustness to critical propagation conditions, ergonomics, and simplicity of use.

14.2.6 MBSon or Raw Audification of Received GSM Signals

The application of Model Based Sonification (MBSon) to DF outputs instead of PMSon as considered in Chapter 8 may well result in efficient designs. A paradigm that would allow users to perceive the complete spatial or frequency power spectrum of the received GSM signal could potentially lead to improved situational awareness. It is quite possible

[3] Different timbres and spatial positions where used for each tone to simplify selective listening.
[4] Eventually employing the auditory zoom paradigms discussed in [155].

that experienced listeners could e.g. learn to perceive spectrum variations related to certain paths or material reflections, as warfare specialist sonar submariners learn to interpret "sonobuoy" information.

In the early stages of this research, a study on raw audification of the received GSM signal was considered, discarded in favor of a more reliable solution. This study came from consideration of the benefits of individualized HRTFs (see Chapter 7), along with the observation that GSM and audio signals have comparable wavelengths ($\lambda_{\text{GSM900MHz}} \approx \lambda_{\text{pure tone of 1 kHz}}$). As HRTFs are spatial filters, function of said wavelength, a set of ears/head/torso that would interact with electromagnetic waves as humans do with acoustic waves could produce a set of "RF HRTFs" with a GSM input equivalent to the one measured with a 1 kHz sound. A simple heterodyne receiver, applying a straightforward band transposition on the received GSM spectrum would then allow one to basically hear the incoming wave as one would hear a sound. Note that even if covering said ears/head/torso with a material that would validate this assumption, one of the remaining problem would be that locating a 1 kHz pure tone is hard, even in real life. In that direction, one could take advantage of the TDMA modulation used in the GSM protocol to produce repeated 1 kHz bursts, easier to localize than a continuous pure tone[5]. Fast shifts with GSM1800 could also be employed to provide a bi-tone listening. One major issue is that, while the wavelengths are comparable, propagation speeds are not, with the speed of sound being 10^6 times slower than the speed of electromagnetic waves. This difference would translate to time-of-arrival differences between left and right ears for sound sources not on the median plane of the listener, e.g. an interaural time difference of approx. 500 µs for the acoustical wave to 500 ps in the electromagnetic domain for a signal DOA on the axis defined by listener's ears.

14.3 Publications and Diffusions

The work presented in this manuscript led to the publication of several conference proceedings, presented and discussed during international conferences. The investigations and studies presented in Chapters 2, 3, and 8 were gathered in a patent describing a cellphone-based localization procedure using an automatic DF combined with binaural rendering. Chapter 9 has been submitted (early 2015) as a journal paper to the International Journal on Human-Computer Studies (Elsevier), for a special issue on interactive sonification. Finally, I had and still have the pleasure of contributing to the development of BlenderVR (previously referred to as BlenderCAVE), the adaptation of the 3D creation content software Blender, used as an authoring tool and rendering engine for virtual reality.

[5] see literature review in Chapter 7.

14.3.1 Journal Article

D. Poirier-Quinot, G. Parseihian, and B. F.G. Katz, "Reduction of perceived instabilities in Parameter Mapping Sonification: application to the real-time exploration of a noisy stream of data" submitted 22–Jan–2015, IJHCS.

14.3.2 Conference Proceedings

D. Poirier-Quinot and B. F.G. Katz, "BlenderVR: Open-Source framework for interactive and immersive VR" accepted for the IEEE VR Conference, March 2015.

D. Poirier-Quinot and B. F.G. Katz, "CAVE-based virtual prototyping of an audio radiogoniometer: Ecological validity assessment" in the 20th International Conference on Auditory Display (ICAD), (NYC), pp. 1-8, ICAD, June 2014. http://hdl.handle.net/1853/52064.

D. Poirier-Quinot, D. Touraine, and B. F.G. Katz, "BlenderCAVE: A multimodal scene graph editor for Virtual Reality" in the 19th International Conference on Auditory Display (ICAD), (Lodz), pp. 223-230, ICAD, Oct 2013. http://hdl.handle.net/1853/51672.

D. Poirier-Quinot, P. Duvaut, L. Girardeau, and B. F.G. Katz, "3D Head-mounted antenna array architecture optimization based on the Fisher Information Matrix" in the International Conference on Ultra Modern Telecommunications (ICUMT), (St. Petersburg), pp. 135-142, Oct. 2012. https://hal.archives-ouvertes.fr/LIMSI/hal-01108665v1.

14.3.3 Patent

D. Poirier-Quinot, F. Cibaud, L. Girardeau, and B. F. Katz,"Procédé d'aide à la localisation d'une personne en détresse grâce à son téléphone portable" (patent pending). FR 14/59915 (submitted 14-oct-2014).

14.3.4 Software

BlenderVR, http://blendervr.limsi.fr

Appendix A

BlenderVR: Open-Source Framework for Interactive and Immersive VR

This chapter succinctly presents BlenderVR, an open-source project framework for interactive and immersive applications based on an extension of the Blender Game Engine to Virtual Reality applications. BlenderVR is a generalization of the BlenderCAVE project, accounting for alternate platforms (e.g., HMD, video-walls). The goal is to provide a flexible and easy to use framework for the creation of Virtual Reality (VR) applications for various platforms, making use of the existing power of the BGE's graphics rendering and physics engine. Compatible with 3 major Operating Systems, BlenderVR has been developed by VR researchers with support from the Blender Community. BlenderVR currently handles multi-screen/multi-user tracked stereoscopic rendering through efficient low-level master/slave synchronization process with multimodal interactions via OSC and VRPN protocols.

A.1 Introduction

We present here the current state of development of the BlenderVR project[1]. BlenderVR is a generalization of the previous BlenderCAVE and BlenderCAVE3D-s projects [3, 263]. BlenderVR is principally a scene graph editor based on the well established Open-source Blender and Blender Game Engine (BGE) [264] software. Blender is a multi-platform open source 3D creation content software[2] with enough functionalities to create photorealistic pictures, high quality animations and, most of all, video games. Blender based games use the BGE real-time rendering engine which handles a multitude of physical interactions through the implemented Bullet Physics Library while general game logic may be defined through blocks and/or embedded python scripts. Gathering users for more than a decade, Blender now boasts a large support community, a dedicated professional network and several scene repositories where one may find plenty of reusable material.

[1] http://BlenderVR.limsi.fr & http://github.com/BlenderVR

[2] http://www.blender.org

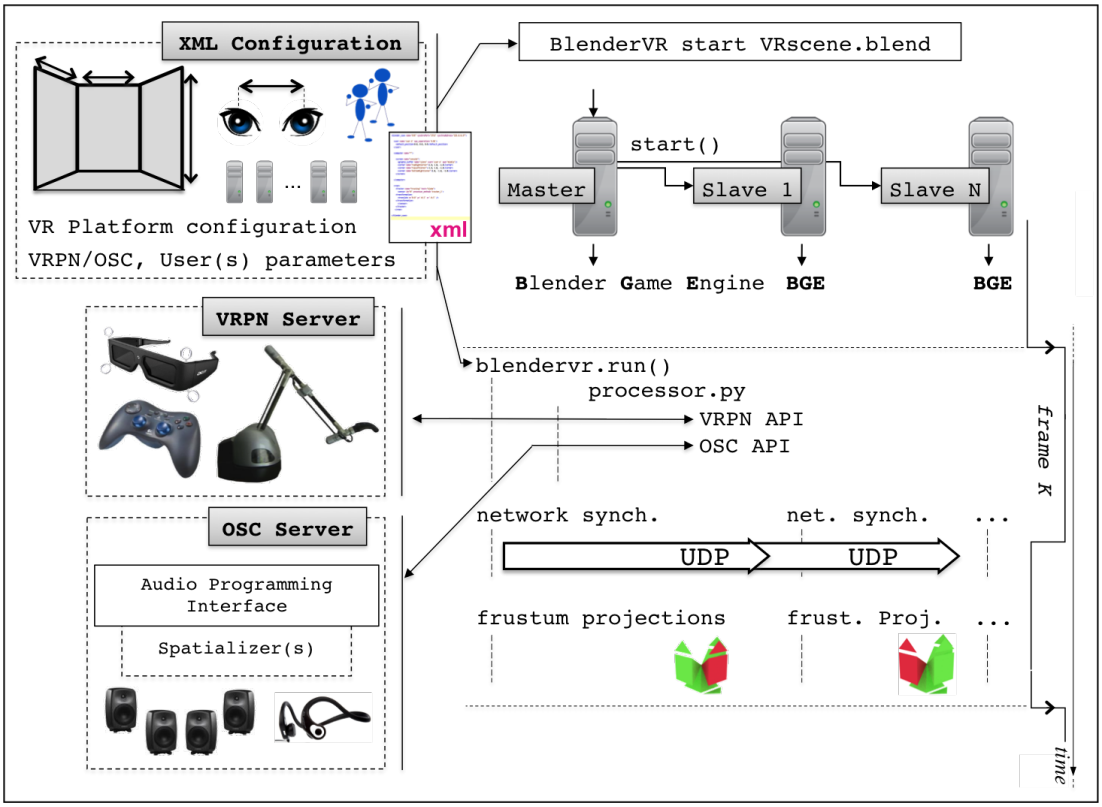


Figure A.1: BlenderVR architecture graph showing (right) the 3 main processes integrated in the BGE frame based sequencer. *network synchronization* process maintains consistency between master/slave rendering nodes. *frustum projection* handles tracked adaptive stereoscopic rendering calculations.

A.2 Main Features

BlenderVR, a patched version of the core Blender software, adds additional functionalities key to VR applications, while benefiting from the basic BGE architecture, interface, and user community. The BlenderVR code project is accessible via GitHub[1] and collaborators are welcome. The project website hosts documentation, executable downloads, and examples.

The core BlenderVR modifications to the BGE master consist in the addition of a `prerender` method prior to the basic `predraw`, the ability to redefine the projection matrix, and the possibility to redefine the aspect-ratio from a python script (allowing easier portability of scenes between architectures). BlenderVR routines, as well as most of the BlenderVR patched modifications, have been implemented so as to be as transparent as possible regarding BGE native processes in the hope that they will soon

be integrated in the official Blender master, eliminating the need for patching.

A.2.1 Master/Slave Synchronization

The synchronisation of multiple graphic nodes is necessary for assuring continuity and coherence of the scene in situations with multiple displays (e.g., CAVE, videowall) or rendering instances (e.g., mutliple HMD). Master/Slave synchronization is carried out at each frame, in the `prerender` method. Synchronization, executed via a python script, inspects every object in the scene to see if it has changed. If so, the update information is passed from the Master to the Slave nodes before rendering. TCP multi-unicast is used for synchronization and acknowledgment messages. By default, BlenderVR synchronizes every object in the scene amongst the rendering nodes. BlenderVR also supports synchronization, white and black lists, forced synchronization without checks or no synchronization or checks for specific objects, thereby reducing unnecessary calculation overhead in complex scene rendering.

Already evaluated for CAVE-type configurations, current development concerns collaborative work with combined HMD and CAVE or multiple HMD devices on shared or remote sites. The importance of synchronization of scene graphs is crucial for such applications and BlenderVR should be well suited to this situation.

A.2.2 Adaptive Stereoscopy for Large-Screen Projections

Adaptive stereoscopy is required for Workbench, projected Walls or CAVE-like systems. This rendering mainly consists in changes of coordinate systems and projection operations. The resulting modification in the projection matrix is necessary because the user is, more often than not, away from the center of the screen and thereby requires non-symmetric projection matrices. This is updated to take into account user's current head position, orientation, and eye separation. Such variable frustum projections have been implemented so as to be computed locally on each graphical node before rendering in the `prerender` process. This implementation is generic enough to manage multi-stereoscopy, which allows for several tracked users to each have correct depth perception.

A.2.3 External Message Processing

While stereoscopic rendering and synchronization are supposedly basic VR features, message exchange with external software represents a cornerstone in any scene graph editor as it largely impacts end users in their scene developments. BlenderVR to external (and vice versa) user defined interactions are conventionally implemented and collected in a python script attached to the VR scene. Incoming messages are often dedicated to user integration or interaction in the VR scene, from position/orientation

coordinates to various controller states, and typically employ the VRPN protocol. For this, BlenderVR implements the *Processor* python Class which integrates dedicated OSC and VRPN APIs to ease message processing for access to controllers and user parameters.

Most outgoing messages are currently intended to add sound to the scene, i.e., update scene graph object properties in the Sound Rendering Engine (SRE) using the well known Open Sound Control (OSC) protocol. The SRE here refers to the system gathering audio object definitions (sound source, position, etc.) and sound rendering methods in an Audio Programming Environment (e.g., Max/MSP, PureData, etc.) plus eventual loudspeaker/headset outputs. The overall system, combining BlenderVR for visual rendering and the SRE for spatially distributed sound sources resembles the framework based on ViSTA and its audio manager introduced in [265]. The OSC API encapsulates three main Classes of messages: Global, Object, and User.

A fourth class of message manages the routing of Objects to multiple spatialization engines (i.e., Users), such that certain Objects can be rendered for selected Users if desired, and is thus named the *ObjectUser* Class. These messages define which engine handles a given set of objects in the scene graph.

A.3 Example Installations

Scene graph edition in BlenderVR is essentially based on Blender creation and animation tools, while audio rendering related events are written in the Processor script. The overall process being architecture independent, most scene development can be carried out on standard laptops, before being ported to the actual VR systems. Once installed, basic import of a Blender scene (without tracking nor audio rendering) is almost transparent.

BlenderVR has been developed on two VR platforms: SMART-I2 [266] and EVE[3]. Portability of developed scenes between such different VR architectures is the driving goal behind the project.

The SMART-I2 implements passive adaptive stereoscopic rendering through a pair of front-projected rigid screens which also serve as a 24 channel loudspeaker array offering horizontal Wave Field Synthesis audio rendering. One computer handles graphics, with 4 projectors for 2 screens, running Ubuntu Precise on an Intel Core 2 2.66 GHz Quad Q9400, 4 GB of RAM, and 2 NVIDIA GeForce GTX 470. Audio is handled separately.

The EVE system offers a multi-user/multi-sensorimotor environment with adaptative double-stereoscopic rendering, 15 loudspeakers and 2 RF modules for wireless audio input and individual binaural renderings. EVE comprises 4 rear-projected screens cou-

[3] http://www.limsi.fr/venise/EVEsystem

pled to 7 projectors providing $\sim 60m^2$ of high definition projection space. The system comprises 8 i7 computers (1/projector + 1 monitoring console) with Ubuntu Precise, 12 GB of RAM, and Quadro 6000. Audio and haptics are handled separately.

A.4 Example Applications

In a study investigating paradigms for 6DoF navigation in immersive virtual worlds, [267] compared joystick-based input devices and steering metaphors based on movements of the user's body, e.g., head-controlled paradigms. BlenderVR served as the software platform, with various input devices controlling the user's flight through a series of navigational tasks. While the virtual world was static, the selection of various test configurations and generation of experimental logs were all achieved within BlenderVR.

To evaluate different designs for a user audio guidance system for rescue workers, [2] implemented virtual prototypes in an attempt to provide ecologically valid test conditions. The BlenderVR based virtual prototypes employed various input devices (e.g., 6 DoF tracking, Wii Balance Board, Wiimote) along with OSC communication with MaxMSP to generate interactive audio feedback sonification based on geometrical data from the BlenderVR scene.

Real-time animation of virtual avatars based on motion capture within the BlenderVR framework was investigated by [268]. A Python module was developed to control avatars using an ART MoCap tracking system, common in VR applications.

Focusing on collaborative situations, [269] investigated navigation in a multi-stereoscopic immersive system (several users sharing the same restricted workspace). In this context, a proper navigation paradigm should provide users both efficient control of virtual navigation and a guaranty of users' safety in the real workspace relative to the display system and between users.

A.5 Conclusion

Due to the lack of affordable scene creation software dedicated to Virtual Reality in augmented environments, the BlenderVR project was created, implementing a new scene graph editor based on the popular Blender Game Engine. To adapt the BGE to VR, BlenderVR includes multi-user tracked stereoscopic rendering, master-slave synchronization, and OSC and VRPN communication methods for external applications. Compatible with Windows, Linux, and MacOS, BlenderVR is released as open source. BlenderVR is aimed at those in need of a straightforward scene development solution for multimodal VR creation.

Bibliography

[1] D. Poirier-Quinot, P. Duvaut, L. Girardeau, and B. F.G.Katz, "3D head-mounted antenna array architecture optimization based on the Fisher Information Matrix," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2012 4th International Congress on*, pp. 135–142, IEEE, 2012. (Cited on pages and 42.)

[2] D. Poirier-Quinot and B. F. Katz, "CAVE-based virtual prototyping of an audio radiogoniometer: Ecological validity assessment," in *Proceedings of the International Conference on Auditory Display (ICAD)*, (NYC), pp. 1–8, June 2014. (Cited on pages , 140 and 171.)

[3] D. Poirier-Quinot, D. Touraine, and B. F. Katz, "BlenderCAVE: A multimodal scene graph editor for virtual reality," in *Proceedings of the International Conference on Auditory Display (ICAD)*, (Lodz), pp. 223–230, Oct 2013. (Cited on pages , 115, 143 and 167.)

[4] Q. Hamp, R. Zhang, L. Chen, O. Gorgis, T. Ostertag, and J. Pavlina, "Results from german research project I-LOV," in *Proceedings of the 1st International Conference on Wireless Technologies for Humanitarian Relief*, pp. 249–258, ACM, 2011. (Cited on page 1.)

[5] SGL, "Second Generation Locator for Urban Search and Rescue Operations". URL: `http://www.sgl-eu.org` [accessed 10/27/2014]. (Cited on page 1.)

[6] Profitex, "Advanced Protective Firefighter Equipment". URL: `https://www.project-profitex.eu` [accessed 10/27/2014]. (Cited on page 1.)

[7] G. de Cubber, D. Doroftei, Y. Baudoin, D. Serrano, K. Chintamani, R. Sabino, and S. Ourevitch, "Icarus: An eu-fp7 project providing unmanned search and rescue tools," in *IROS Workshop on Robots and Sensors integration in future rescue INformation system (ROSIN)*, 2012. (Cited on page 1.)

[8] J. Wickman and A. Hakansson, "Localization within GSM," in *Royal Institute of Navigation Conference: land navigation and information systems*, pp. 7–16, 1990. (Cited on page 1.)

[9] S. Shek, "Next-generation location-based services for mobile devices," in *Leading Edge Forum, Computer Science Corporation*, pp. 1–66, 2010. (Cited on page 1.)

[10] F. C. Commission *et al.*, "Revision of the commission's rules to ensure compatibility with enhanced 911 emergency calling systems," *Report and Order and Further Notice of Proposed Rulemaking, Tech. Rep. CC Docket*, pp. 94–102, 1996. (Cited on page 1.)

[11] B. Sanou, "The world in 2014: ICT facts and figures," *International Telecommunication Union*, 2014. (Cited on pages 1 and 6.)

[12] K. G. Jansky, "Radio waves from outside the solar system," *Nature*, vol. 132, pp. 66–66, 1933. (Cited on page 2.)

[13] A.-M. Roxin, J. Gaber, M. Wack, A. N. S. Moh, *et al.*, "Survey of wireless geolocation techniques," in *IEEE Globecom Workshops*, pp. 1–9, 2007. (Cited on page 2.)

[14] Proximus, "Mirage, Real-Time GSM/UMTS mobile terminals tracking platform". URL: http://www.proximus.com.ua/Mirage_location_detector.html [accessed 10/28/2014]. (Cited on page 2.)

[15] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *Communications Magazine, IEEE*, vol. 48, no. 9, pp. 140–150, 2010. (Cited on page 2.)

[16] B. Edgerly and J. Hereford, "Digital transceiving systems: The next generation of avalanche beacons," in *Proceedings of the International Snow Science Workshop*, pp. 120–127, 1998. (Cited on page 2.)

[17] P. Morgand, A. Ferreol, R. Sarkis, C. Craeye, and C. Oestges, "Detection and location of people in emergency situations through their PMR or GSM/UMTS phones," in *Wireless Technology Conference (EuWIT), 2010 European*, pp. 185–188, IEEE, 2010. (Cited on pages 2 and 60.)

[18] Assasar, "ASSA Search and Rescue". URL: http://assasar.com [accessed 10/28/2014]. (Cited on page 3.)

[19] S. Zorn, R. Rose, A. Goetz, and R. Weigel, "A novel technique for mobile phone localization for search and rescue applications," in *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on*, pp. 1–4, IEEE, 2010. (Cited on pages 3, 10, 121, 133, 134, 156 and 164.)

[20] Raven-Research, "Wideband Hand Held Radio DF". URL: http://www.electronicsarena.co.uk/companies/raven-research/products/wideband-hand-held-radio-df [accessed 10/28/2014]. (Cited on page 3.)

[21] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997. (Cited on pages 3, 62 and 65.)

[22] P. B. Meijer, "An experimental system for auditory image representations," *Biomedical Engineering, IEEE Transactions on*, vol. 39, no. 2, pp. 112–121, 1992. (Cited on page 3.)

[23] J. L. González-Mora, A. Rodriguez-Hernandez, E. Burunat, F. Martin, and M. Castellano, "Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people," in *Information and Communication Technologies (ICTTA)*, vol. 1, pp. 837–842, IEEE, 2006. (Cited on page 3.)

[24] 3GPP, "The mobile broadband standard". URL: `http://www.3gpp.org/specifications` [accessed 10/18/2014]. (Cited on page 6.)

[25] M. Mouly, M.-B. Pautet, and T. Foreword By-Haug, *The GSM system for mobile communications*. Telecom Publishing, 1992. (Cited on page 6.)

[26] G. Heine and M. Horrer, *GSM networks: protocols, terminology, and implementation*. Artech House, Inc., 1999. (Cited on page 6.)

[27] P. W. Baier, "A critical review of CDMA," in *46th Vehicular Technology Conference on Mobile Technology for the Human Race*, vol. 1, pp. 6–10, IEEE, 1996. (Cited on page 8.)

[28] K. Murota and K. Hirade, "GMSK modulation for digital mobile radio telephony," *IEEE Transactions on Communications*, vol. 29, no. 7, pp. 1044–1050, 1981. (Cited on page 8.)

[29] ETSI, TS, "Digital cellular telecommunications system (Phase 2+); Modulation," in *Technical Specification, 3GPP TS 45.004 version 6.0. 0 Release 6, 145 004 V6.0.0*, 2005. (Cited on page 9.)

[30] E. Barkan, E. Biham, and N. Keller, "Instant ciphertext-only cryptanalysis of GSM encrypted communication," *Journal of Cryptology*, vol. 21, no. 3, pp. 392–429, 2008. (Cited on page 9.)

[31] F. Van Den Broek, "Eavesdropping on GSM: state-of-affairs," in *Benelux Workshop on Information and System Security (WISSec)*, pp. 1–16, 2011. (Cited on page 10.)

[32] K. Boman, G. Horn, P. Howard, and V. Niemi, "UMTS security," *Electronics & Communication Engineering Journal*, vol. 14, no. 5, pp. 191–204, 2002. (Cited on page 10.)

[33] H. Hopkins and B. Pressey, "Current direction-finding practice," *Proceedings of the IEEE–Part B: Radio and Electronic Engineering*, vol. 105, no. 9, pp. 307–316, 1958. (Cited on page 13.)

[34] J. I. Allsop and G. J. Mc Morrow, "Avalanche victim locating apparatus", US Patent, US 4850031 A, July 1989. (Cited on page 13.)

[35] T. E. Tuncer and B. Friedlander, *Classical and modern direction-of-arrival estimation*. Access Online via Elsevier, 2009. (Cited on pages 13, 30 and 35.)

[36] S. Chandran, *Advances in Direction-of-Arrival Estimation*. Artech House, Inc., 2005. (Cited on pages 13 and 29.)

[37] Z. Chen, G. Gokeda, and Y. Yu, *Introduction to Direction-of-arrival Estimation*. Artech House, 2010. (Cited on pages 13 and 25.)

[38] S. Arberet, R. Gribonval, and F. Bimbot, "A robust method to count and locate audio sources in a multichannel underdetermined mixture," *Signal Processing, IEEE Transactions on*, vol. 58, no. 1, pp. 121–133, 2010. (Cited on page 14.)

[39] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *Signal Processing Magazine, IEEE*, vol. 13, no. 4, pp. 67–94, 1996. (Cited on pages 14, 29 and 30.)

[40] M. S. Bartlett, "Smoothing periodograms from time series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, 1948. (Cited on page 19.)

[41] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969. (Cited on page 20.)

[42] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, 2003. (Cited on page 20.)

[43] G. Bienvenu and L. Kopp, "Adaptivity to background noise spatial coherence for high resolution passive methods," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).*, vol. 5, pp. 307–310, April 1980. (Cited on page 22.)

[44] R. O. Schmidt, *A signal subspace approach to multiple emitter location spectral estimation*. PhD thesis, Ph. D. Thesis, Stanford University, 1981. (Cited on page 22.)

[45] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276–280, mar 1986. (Cited on page 22.)

[46] R. Roy and T. Kailath, "ESPRIT estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing (ASSP)*, vol. 37, pp. 984 –995, jul 1989. (Cited on page 24.)

[47] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 8, pp. 336–339, 1983. (Cited on page 25.)

[48] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing,*, vol. 34, no. 5, pp. 1081–1089, 1986. (Cited on page 25.)

[49] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," in *IEEE Radar and Signal Processing*, vol. 137, pp. 19–26, IET, 1990. (Cited on page 25.)

[50] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing (ASSP)*, vol. 38, no. 7, pp. 1132–1143, 1990. (Cited on page 25.)

[51] H. L. van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. John Wiley & Sons, 2004. (Cited on pages 27 and 29.)

[52] W. Tao, Y. Li-sheng, L. Jian-mei, and Y. Shi-zhong, "A modified MUSIC to estimate DOA of the coherent narrowband sources based on UCA," in *International Conference on Communication Technology (ICCT)*, pp. 1–4, IEEE, 2006. (Cited on pages 28 and 54.)

[53] H. Minghao, Y. Yixin, and Z. Xianda, "UCA-ESPRIT algorithm for 2-D angle estimation," in *5th International Conference on Signal Processing Proceedings (WCCC-ICSP)*, vol. 1, pp. 437–440, IEEE, 2000. (Cited on pages 28 and 54.)

[54] K. Maheswara Reddy and V. Reddy, "Analysis of spatial smoothing with uniform circular arrays," *Signal Processing, IEEE Transactions on*, vol. 47, no. 6, pp. 1726–1730, 1999. (Cited on pages 28 and 54.)

[55] R. Goossens, H. Rogier, and S. Werbrouck, "UCA Root-MUSIC with sparse uniform circular arrays," *Signal Processing, IEEE Transactions on*, vol. 56, no. 8, pp. 4095–4099, 2008. (Cited on pages 28 and 59.)

[56] M. Schoor and B. Yang, "Subspace based DOA estimation in the presence of correlated signals and model errors," in *International Conference on Acoustics,*

Speech and Signal Processing (ICASSP), pp. 2161–2164, IEEE, 2009. (Cited on page 29.)

[57] W. Li, W. Yao, and P. J. Duffett-Smith, "Comparative study of joint TOA/DOA estimation techniques for mobile positioning applications," in *Proceedings of the 6th Consumer Communications and Networking Conference*, pp. 1–5, IEEE, 2009. (Cited on page 29.)

[58] F. Harabi, A. Gharsallah, and S. Marcos, "Three-dimensional antennas array for the estimation of direction of arrival," *IET microwaves, antennas & propagation*, vol. 3, no. 5, pp. 843–849, 2009. (Cited on page 29.)

[59] Z.-f. Ye, Y.-f. Zhang, and C. Liu, "Direction-of-arrival estimation for uncorrelated and coherent signals with fewer sensors," *Microwaves, Antennas & Propagation, IET*, vol. 3, no. 3, pp. 473–482, 2009. (Cited on page 29.)

[60] A. Kisliansky, R. Shavit, and J. Tabrikian, "Direction of arrival estimation in the presence of noise coupling in antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 7, pp. 1940–1947, 2007. (Cited on page 29.)

[61] P. Yang, F. Yang, and Z.-P. Nie, "Doa estimation with sub-array divided technique and interporlated ESPRIT algorithm on a cylindrical conformal array antenna," *Progress In Electromagnetics Research*, vol. 103, pp. 201–216, 2010. (Cited on page 30.)

[62] S. Zekavat, A. Kolbus, X. Yang, Z. Wang, J. Pourrostam, and M. Pourkhaatoun, "A novel implementation of DOA estimation for node localization on software defined radios: achieving high performance with low complexity," in *International Conference on Signal Processing and Communications (ICSPC)*, pp. 983–986, IEEE, 2007. (Cited on page 30.)

[63] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, no. 3, pp. 81–91, 1945. (Cited on page 33.)

[64] R. A. Fisher, "The logic of inductive inference," *Journal of the Royal Statistical Society*, vol. 98, no. 1, pp. 39–82, 1935. (Cited on pages 33 and 45.)

[65] A. Dogandzic and A. Nehorai, "Cramer-Rao bounds for estimating range, velocity, and direction with an active array," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1122–1137, June 2001. (Cited on pages 35 and 50.)

[66] K. Binder and D. Heermann, *Monte Carlo simulation in statistical physics: an introduction*. Springer, 2010. (Cited on page 36.)

[67] B. Porat and B. Friedlander, "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," in *IEEE Trans. Acoustics, Speech, and Signal Processing (ASSP)*, vol. 36, pp. 532–544, April 1988. (Cited on pages 37 and 46.)

[68] W. Mohr, "The WINNER (Wireless World Initiative New Radio) project–development of a radio interface for systems beyond 3G," *International Journal of Wireless Information Networks*, vol. 14, no. 2, pp. 67–78, 2007. (Cited on page 38.)

[69] J. Zhang, "Review of wideband mimo channel measurement and modeling for imt-advanced systems," *Chinese Science Bulletin*, vol. 57, no. 19, pp. 2387–2400, 2012. (Cited on page 38.)

[70] P. Kyösti, J. Meinilä, L. Hentilä, X. Zhao, T. Jämsä, C. Schneider, M. Narandzic, M. Milojevic, A. Hong, J. Ylitalo, *et al.*, "WINNER II channel models (d1. 1.2 v1. 1)," in *Radio Technologies and Concepts for IMT-Advanced*, vol. 1, ch. 3, John Wiley & Sons, 2007. (Cited on page 38.)

[71] G. Del Galdo and M. Haardt, "Ilmprop: A flexible geometry-based simulation environment for multi-user MIMO communications," in *COST 273 Proj. Rep. TD (03)*, vol. 188, pp. 1–10, 2003. (Cited on pages 39, 110, 144 and 162.)

[72] A. Manikas, A. Alexiou, and H. Karimi, "Comparison of the ultimate direction-finding capabilities of a number of planar array geometries," *Radar, Sonar and Navigation, IEE Proceedings*, vol. 144, pp. 321–329, Dec 1997. (Cited on page 43.)

[73] J. Verhaevert, E. Van Lil, and A. Van de Capelle, "Uniform spherical distributions for adaptive array applications," in *Vehicular Technology Conference (VTC) IEEE VTS 53rd*, vol. 1, pp. 98 –102, 2001. (Cited on page 43.)

[74] P. Stoica, E. Larsson, and A. Gershman, "The stochastic CRB for array processing: a textbook derivation," *Signal Processing Letters, IEEE*, vol. 8, pp. 148 –150, may 2001. (Cited on page 43.)

[75] A. Ferréol and P. Chevalier, "High resolution direction finding: From performance toward antenna array optimization, the mono-source case," in *Proceedings of European signal processing conference (EUSIPCO)*, pp. 1973–1977, 2009. (Cited on page 43.)

[76] H. Gazzah and S. Marcos, "Cramer-Rao bounds for antenna array design," *IEEE Transactions on Signal Processing*, vol. 54, pp. 336–345, January 2006. (Cited on pages 44 and 50.)

[77] C. Proukakis and A. Manikas, "Study of ambiguities of linear arrays," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 549–552, April 1994. (Cited on page 46.)

[78] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 1, pp. 8–15, 1989. (Cited on pages 53 and 59.)

[79] W. Du and R. L. Kirlin, "Improved spatial smoothing techniques for DOA estimation of coherent signals," *Signal Processing, IEEE Transactions on*, vol. 39, no. 5, pp. 1208–1210, 1991. (Cited on pages 53 and 59.)

[80] C. P. Mathews and M. D. Zoltowski, "2-D angle estimation with uniform circular arrays," *Signal Processing, IEEE Transactions on*, vol. 42, no. 9, pp. 2395–2407, 1994. (Cited on page 54.)

[81] R. Sarkis, C. Craeye, A. Férréol, and P. Morgand, "Design of triple band antenna array for GSM/DCS/UMTS handset localization," in *Antennas and Propagation, 2009. EuCAP 2009. 3rd European Conference on*, pp. 3051–3054, IEEE, 2009. (Cited on pages 59, 60, 125, 156 and 164.)

[82] R. Goossens and H. Rogier, "Closed-form 2d angle estimation with a spherical array via spherical phase mode excitation and ESPRIT," in *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP)*, pp. 2321–2324, 2008. (Cited on page 59.)

[83] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. H. Flowers, N. Miner, and J. Neuhoff, "Sonification report: Status of the field and research agenda," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–30, 1999. (Cited on pages 62 and 63.)

[84] H. Fastl and E. Zwicker, *Psychoacoustics: facts and models*. Springer, 2001. (Cited on pages 62, 63 and 64.)

[85] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag Berlin, 2011. (Cited on page 62.)

[86] B. C. Moore, *An introduction to the psychology of hearing*. Brill, 2012. (Cited on page 62.)

[87] S. C. Peres, V. Best, D. Brock, C. Frauenberger, T. Hermann, J. G. Neuhoff, L. Nickerson, B. Shinn-Cunningham, and A. Stockman, "Auditory interfaces," *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*, pp. 147–195, 2008. (Cited on page 62.)

[88] A. S. Bregman, *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994. (Cited on page 63.)

[89] J. Anderson and P. Sanderson, "Designing sonification for effective attentional control in complex work domains," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 48, pp. 1818–1822, SAGE Publications, 2004. (Cited on pages 63 and 71.)

[90] E. Jovanov, D. Starcevic, D. Karron, K. Wegner, and V. Radivojevic, "Acoustic rendering as support for sustained attention during biomedical procedures," in *Proceedings of the International Conference on Auditory Display (ICAD)*, vol. 98, pp. 1–5, 1998. (Cited on page 63.)

[91] J. Verschuure and A. Van Meeteren, "The effect of intensity on pitch," *Acta Acustica united with Acustica*, vol. 32, no. 1, pp. 33–44, 1975. (Cited on page 63.)

[92] H. G. Tekman, "Interactions of perceived intensity, duration, and pitch in pure tone sequences," *Music Perception*, pp. 281–294, 1997. (Cited on page 63.)

[93] S. S. Stevens, J. Volkmann, and E. B. Newman, "A scale for the measurement of the psychological magnitude pitch," *The Journal of the Acoustical Society of America*, vol. 8, no. 3, pp. 185–190, 1937. (Cited on pages 63, 96 and 164.)

[94] S. S. Stevens, "A scale for the measurement of a psychological magnitude: loudness.," *Psychological Review*, vol. 43, no. 5, p. 405, 1936. (Cited on page 63.)

[95] D. O'Shaughnessy, *Speech communication: human and machine*. Addison-Wesley, 1987. (Cited on page 64.)

[96] Acoustics International Organization for Standardization, "Normal equal-loudness level contours - ISO 226", 2003. (Cited on pages 64 and 96.)

[97] J. D. McAuley, "Tempo and rhythm," in *Music perception*, pp. 165–199, Springer, 2010. (Cited on pages 64 and 93.)

[98] M. R. Jones, "Time, our lost dimension: toward a new theory of perception, attention, and memory.," *Psychological review*, vol. 83, no. 5, pp. 323–355, 1976. (Cited on page 64.)

[99] J. Krimphoff, S. McAdams, and S. Winsberg, "Caractérisation du timbre des sons complexes, II. Analyses acoustiques et quantification psychophysique," *Le Journal de Physique*, vol. 4, no. C5, pp. 5–625, 1994. (Cited on page 65.)

[100] D. R. Begault, *3-D sound for virtual reality and multimedia*. Citeseer, 2000. (Cited on page 65.)

[101] P. M. Hofman, J. G. Van Riswick, and A. J. Van Opstal, "Relearning sound localization with new ears," *Nature neuroscience*, vol. 1, no. 5, pp. 417–421, 1998. (Cited on pages 65 and 79.)

[102] P. Zahorik, P. Bangayan, V. Sundareswaran, K. Wang, and C. Tam, "Perceptual recalibration in human sound localization: Learning to remediate front-back reversals," *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 343–359, 2006. (Cited on page 65.)

[103] S. S. Stevens and E. B. Newman, "The localization of actual sources of sound," *The American Journal of Psychology*, pp. 297–306, 1936. (Cited on page 65.)

[104] D. R. Begault, "Overview of spatial hearing," in *3-D sound for virtual reality and multimedia*, ch. 2, Citeseer, 2000. (Cited on page 66.)

[105] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. II: Psychophysical validation," *The Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 868–878, 1989. (Cited on page 66.)

[106] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using non-individualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993. (Cited on pages 66 and 79.)

[107] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 904–916, 2001. (Cited on page 66.)

[108] E. Knudsen and M. Brainard, "Creating a unified representation of visual and auditory space in the brain," *Annual review of neuroscience*, vol. 18, no. 1, pp. 19–43, 1995. (Cited on page 66.)

[109] S. Carlile, "The plastic ear and perceptual relearning in auditory spatial perception," *Frontiers in neuroscience*, vol. 8, 2014. (Cited on page 66.)

[110] S. A. Brewster, P. C. Wright, and A. D. Edwards, "A detailed investigation into the effectiveness of earcons," in *Proceedings of Santa Fe Institute Studies in the Sciences of Complexity*, vol. 18, pp. 471–471, 1994. (Cited on page 67.)

[111] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human–Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989. (Cited on page 67.)

[112] W. Gaver, "The sonic finder: An interface that uses auditory icons. the use of non-speech audio at the interface," in *Human Computer Interaction*, vol. 4, pp. 67–94, 1989. (Cited on page 67.)

[113] T. Dingler, J. Lindsay, and B. N. Walker, "Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–6, 2008. (Cited on page 67.)

[114] B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in *Proceedings of the International Conference on Auditory Display (ICAD)*, (UK), pp. 63–68, 2006. (Cited on page 67.)

[115] B. N. Walker and G. Kramer, "Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making," *Ecological psychoacoustics*, pp. 150–175, 2004. (Cited on page 67.)

[116] C. A. Giller, A. M. Murro, Y. Park, S. Strickland, and J. R. Smith, "EEG sonification for epilepsy surgery: A clinical work-in progress," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 225–226, 2012. (Cited on page 67.)

[117] F. Dombois, "Auditory seismology on free oscillations, focal mechanisms, explosions and synthetic seismograms," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–4, 2002. (Cited on page 67.)

[118] S. Pauletto and A. Hunt, "A comparison of audio & visual analysis of complex time-series data sets," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 175–181, 2005. (Cited on page 68.)

[119] B. N. Walker, "Magnitude estimation of conceptual data dimensions for use in sonification," *Journal of Experimental Psychology: Applied*, vol. 8, no. 4, pp. 211–221, 2002. (Cited on pages 68 and 95.)

[120] B. N. Walker and G. Kramer, "Mappings and metaphors in auditory displays: An experimental assessment," *ACM Transactions on Applied Perception (TAP)*, vol. 2, no. 4, pp. 407–412, 2005. (Cited on page 68.)

[121] F. Grond and J. Berger, "Parameter mapping sonification," *The Sonification Handbook*, pp. 363–397, 2011. (Cited on pages 68 and 96.)

[122] T. Hermann and H. Ritter, "Listen to your data: Model-based sonification for data analysis," *Advances in intelligent computing and multimedia systems*, vol. 8, pp. 189–194, 1999. (Cited on pages 68 and 72.)

[123] T. Hermann, *Sonification for exploratory data analysis*. PhD thesis, Bielefeld University, Bielefeld, Germany, 2002. (Cited on page 68.)

[124] T. Hermann, P. Meinicke, and H. Ritter, "Principal curve sonification," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–6, Citeseer, 2000. (Cited on pages 68 and 73.)

[125] R. Tünnermann and T. Hermann, "Multi-touch interactions for model-based sonification," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–7, 2009. (Cited on pages 68 and 72.)

[126] S. McAdams and E. Bigand, "Introduction to auditory cognition," *Thinking in sound*, pp. 1–9, 1993. (Cited on page 69.)

[127] D. Kish and H. Bleier, "Echolocation: What it is, and how it can be taught and learned," *California Association of Orientation and Mobility Specialists*, 2000. (Cited on pages 69 and 73.)

[128] V. P., "Sonification for process monitoring," *The Sonification Handbook*, pp. 475–512, 2011. (Cited on page 69.)

[129] T. Hermann, "Taxonomy and definitions for sonification and auditory display," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–8, 2008. (Cited on page 69.)

[130] B. Terri L. and F. John H., "Evaluation of auditory display," *The Sonification Handbook*, pp. 131–164, 2011. (Cited on page 69.)

[131] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey, *Human-Computer Interaction*. Addison-Wesley Longman Ltd., 1994. (Cited on page 69.)

[132] P. Janata and E. Childs, "Marketbuzz: Sonification of real-time financial data.," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–7, 2004. (Cited on pages 69 and 73.)

[133] J. Edworthy, "Does sound help us to work better with machines? a commentary on rauterberg's paper 'about the importance of auditory alarms during the operation of a plant simulator'," *Interacting with Computers*, vol. 10, no. 4, pp. 401–409, 1998. (Cited on page 70.)

[134] G. Parseihian and B. F. Katz, "Morphocons: A new sonification concept based on morphological earcons," *Journal of the Audio Engineering Society*, vol. 60, no. 6, pp. 409–418, 2012. (Cited on pages 70 and 74.)

[135] G. Kramer, "A letter from Greg Kramer: founder of ICAD," *Proceedings of the International Workshop on Interactive Sonification*, pp. 1–2, January 2004. (Cited on page 70.)

[136] D. Lunney and R. C. Morrison, "High technology laboratory aids for visually handicapped chemistry students," *Journal of Chemical Education*, vol. 58, no. 3, pp. 228–231, 1981. (Cited on page 70.)

[137] W. T. Fitch and G. Kramer, "Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system," in *Proceedings of Santa Fe Institute Studies in the Sciences of Complexity*, vol. 18, pp. 307–325, 1994. (Cited on page 70.)

[138] W. W. Gaver, R. B. Smith, and T. O'Shea, "Effective sounds in complex systems: The arkola simulation," in *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pp. 85–90, ACM, 1991. (Cited on page 70.)

[139] T. Hermann and A. Hunt, "The discipline of interactive sonification," in *Proceedings of the Int. Workshop on Interactive Sonification (ISon)*, pp. 1–9, 2004. (Cited on page 70.)

[140] T. Hermann and A. Hunt, "An introduction to interactive sonification," *IEEE Multimedia*, pp. 20–24, 2005. (Cited on page 70.)

[141] B. N. Walker and J. T. Cothran, "Sonification sandbox: A graphical toolkit for auditory graphs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, vol. 3, pp. 161–163, 2003. (Cited on page 71.)

[142] L. M. Brown and S. A. Brewster, "Drawing by ear: Interpreting sonified line graphs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–4, 2003. (Cited on page 71.)

[143] K. Beilharz, "Wireless gesture controllers to affect information sonification," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 105–112, 2005. (Cited on page 71.)

[144] H. Zhao, C. Plaisant, and B. Shneiderman, "I hear the pattern: Interactive sonification of geographical data patterns," in *CHI Extended Abstracts on Human Factors in Computing Systems*, pp. 1905–1908, ACM, 2005. (Cited on page 71.)

[145] A. S. Bregman, "Auditory scene analysis," in *Proceedings of the 7th International Conference on Pattern Recognition*, pp. 168–175, 1984. (Cited on page 71.)

[146] H. J. Song and K. Beilharz, "Aesthetic and auditory enhancements for multi-stream information sonification," in *Proceedings of the International Conference on Digital Interactive Media in Entertainment and Arts*, pp. 224–231, ACM, 2008. (Cited on page 71.)

[147] B. F. Katz and L. Picinali, "Spatial audio applied to research with the blind," *Advances in Sound Localization. InTech*, pp. 225–250, 2011. (Cited on page 71.)

[148] I. Pollack and J. Pickett, "Cocktail party effect," *The Journal of the Acoustical Society of America*, vol. 29, no. 11, pp. 1262–1262, 1957. (Cited on page 71.)

[149] B. Arons, "A review of the cocktail party effect," *Journal of the American Voice I/O Society*, vol. 12, no. 7, pp. 35–50, 1992. (Cited on page 71.)

[150] H. J. Song, K. Beilharz, and D. Cabrera, "Evaluation of spatial presentation in sonification for identifying concurrent audio streams," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 285–292, 2007. (Cited on page 71.)

[151] J. E. Anderson, *Sonification design for complex work domains: Streams, mappings and attention*. PhD thesis, University of Queensland, School of Psychology, 2004. (Cited on page 71.)

[152] M. Watson and P. Sanderson, "Sonification supports eyes-free respiratory monitoring and task time-sharing," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 46, no. 3, pp. 497–517, 2004. (Cited on page 71.)

[153] J. H. Schuett, R. J. Winton, J. M. Batterman, and B. N. Walker, "Auditory weather reports: demonstrating listener comprehension of five concurrent variables," in *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound*, p. 17, ACM, 2014. (Cited on page 71.)

[154] B. Snyder, *Music and memory: An introduction*. MIT press, 2000. (Cited on page 71.)

[155] T. Bouchara, B. Katz, C. Jacquemin, and C. Guastavino, "Audio-visual renderings for multimedia navigation," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 245–252, 2010. (Cited on pages 72 and 164.)

[156] T. Bouchara, C. Jacquemin, and B. F. Katz, "Cueing multimedia search with audiovisual blur," *ACM Transactions on Applied Perception (TAP)*, vol. 10, no. 2, pp. 1–21, 2013. (Cited on page 72.)

[157] J. H. Flowers, "Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions," *Faculty Publications, Department of Psychology*, p. 430, 2005. (Cited on pages 72 and 91.)

[158] T. Stockman, L. V. Nickerson, and G. Hind, "Auditory graphs: A summary of current experience and towards a research agenda," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 420–422, 2005. (Cited on pages 72 and 73.)

[159] S. Wall and S. Brewster, "Feeling what you hear: tactile feedback for navigation of audio graphs," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1123–1132, ACM, 2006. (Cited on page 72.)

[160] R. Ramloll, W. Yu, S. Brewster, B. Riedel, M. Burton, and G. Dimigen, "Constructing sonified haptic line graphs for the blind student: first steps," in *Proceedings of the fourth international ACM conference on Assistive technologies*, pp. 17–25, ACM, 2000. (Cited on page 72.)

[161] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 1, pp. 25–35, 2010. (Cited on page 72.)

[162] B. F. Katz, F. Dramas, G. Parseihian, O. Gutierrez, S. Kammoun, A. Brilhault, L. Brunet, M. Gallay, B. Oriola, M. Auvray, *et al.*, "NAVIG: guidance system for the visually impaired using virtual augmented reality," *Technology and Disability*, vol. 24, no. 2, pp. 163–178, 2012. (Cited on page 72.)

[163] F. W. Youlton, "Portable narrow angle sonar range finder", US Patent, US 5341346 A, August 1994. (Cited on page 72.)

[164] S. Barrass and B. Zehner, "Responsive sonification of well-logs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–8, Citeseer, 2000. (Cited on page 72.)

[165] B. F. Katz, E. Rio, L. Picinali, and O. Warusfel, "The effect of spatialization in a data sonification exploration task," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–7, 2008. (Cited on pages 72 and 91.)

[166] D. R. Smith and B. N. Walker, "Tick-marks, axes, and labels: The effects of adding context to auditory graphs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–6, 2002. (Cited on page 72.)

[167] D. R. Smith and B. N. Walker, "Effects of auditory context cues and training on performance of a point estimation sonification task," *Applied Cognitive Psychology*, vol. 19, no. 8, pp. 1065–1087, 2005. (Cited on page 72.)

[168] B. N. Walker, M. A. Nees, *et al.*, "An agenda for research and development of multimodal graphs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–5, 2005. (Cited on page 72.)

[169] T. Hermann, J. Krause, and H. Ritter, "Real-time control of sonification models with a haptic interface," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 82–86, 2002. (Cited on page 72.)

[170] S. Barrass, "Digital fabrication of acoustic sonifications," *Journal of the Audio Engineering Society*, vol. 60, no. 9, pp. 709–715, 2012. (Cited on page 72.)

[171] S. Barrass, "The hypertension singing bowl: Research through design in acoustic sonification," in *NIME Workshop*, pp. 1–3, 2014. (Cited on page 73.)

[172] T. A. Stroffregen and J. B. Pittenger, "Human echolocation as a basic form of perception and action," *Ecological psychology*, vol. 7, no. 3, pp. 181–216, 1995. (Cited on page 73.)

[173] C. Arias and O. A. Ramos, "Psychoacoustic tests for the study of human echolocation ability," *Applied Acoustics*, vol. 51, no. 4, pp. 399–419, 1997. (Cited on page 73.)

[174] G. Baier and T. Hermann, "The sonification of rhythms in human electroencephalogram," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–5, 2004. (Cited on page 73.)

[175] T. Hinterberger and G. Baier, "POSER: Parametric orchestral sonification of EEG in real-time for the self-regulation of brain states," *IEEE Trans. Multimedia*, vol. 12, pp. 70–76, 2005. (Cited on page 73.)

[176] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H. M. Müller, and S. Weiss, "Sonification for EEG data analysis," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 37–41, 2002. (Cited on page 73.)

[177] S. C. Peres and D. M. Lane, "Sonification of statistical graphs," in *Proceedings of the International Conference on Auditory Display (ICAD)*, (Boston, MA), 2003. (Cited on page 73.)

[178] G. Baier, T. Hermann, O. M. Lara, and M. Müller, "Using sonification to detect weak cross-correlations in coupled excitable systems," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 312–315, 2005. (Cited on page 73.)

[179] M. H. Hansen and B. Rubin, "Babble online: applying statistics and design to sonify the internet," in *Design guidelines for audio presentation of graphs and tables*, pp. 10–15, Citeseer, 2001. (Cited on page 73.)

[180] S. Pauletto and A. Hunt, "Interactive sonification of complex data," *International Journal of Human-Computer Studies*, vol. 67, no. 11, pp. 923–933, 2009. (Cited on page 73.)

[181] L. M. Brown, S. A. Brewster, S. Ramloll, R. Burton, and B. Riedel, "Design guidelines for audio presentation of graphs and tables," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 1–4, 2003. (Cited on pages 73 and 96.)

[182] J. L. Reuss, "Pulse oximeter with signal sonification", US Patent, US 6449501 B1, September 2002. (Cited on page 74.)

[183] I. Wallis, T. Ingalls, T. Rikakis, L. Olsen, Y. Chen, W. Xu, and H. Sundaram, "Real-time sonification of movement for an immersive stroke rehabilitation environment," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 497–503, 2007. (Cited on page 74.)

[184] S. C. Peres and D. M. Lane, "Auditory graphs: The effects of redundant dimensions and divided attention," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 169–174, Citeseer, 2005. (Cited on page 74.)

[185] D. Doukhan and A. Sédès, "CW_binaural: A binaural synthesis external for pure data," in *Puredata International Convention Proceedings*, 2009. (Cited on page 78.)

[186] R. Gilkey and T. R. Anderson, *Binaural and spatial hearing in real and virtual environments*. Psychology Press, 2014. (Cited on page 78.)

[187] F. L. Wightman and D. J. Kistler, "Resolution of front–back ambiguity in spatial hearing by listener and source movement," *The Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2841–2853, 1999. (Cited on page 79.)

[188] B. F. Katz and G. Parseihian, "Perceptually based head-related transfer function database optimization," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. 99–105, 2012. (Cited on page 79.)

[189] G. Parseihian and B. F. Katz, "Rapid head-related transfer function adaptation using a virtual auditory environment," *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2948–2957, 2012. (Cited on page 79.)

[190] A. Lindau, J. Estrella, and S. Weinzierl, "Individualization of dynamic binaural synthesis by real time manipulation of ITD," in *Audio Engineering Society Convention 128*, pp. 1051–1052, 2010. (Cited on page 79.)

[191] B. N. Walker, J. Lindsay, *et al.*, "Navigation performance in a virtual environment with bonephones," in *Proceedings of the International Conference on Auditory Display (ICAD)*, vol. 3, pp. 1–26, 2005. (Cited on page 79.)

[192] J. Hebrank and D. Wright, "Spectral cues used in the localization of sound sources on the median plane," *The Journal of the Acoustical Society of America*, vol. 56, no. 6, pp. 1829–1834, 1974. (Cited on page 88.)

[193] B. N. Walker and D. M. Lane, "Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 90–94, 2001. (Cited on page 88.)

[194] M. S. Mannoor, Z. Jiang, T. James, Y. L. Kong, K. A. Malatesta, W. O. Soboyejo, N. Verma, D. H. Gracias, and M. C. McAlpine, "3D printed bionic ears," *Nano letters*, vol. 13, no. 6, pp. 2634–2639, 2013. (Cited on page 89.)

[195] G. Parseihian, C. Gondre, M. Aramaki, and R. Kronland, "Exploring the usability of sound strategies for guiding task: toward a generalization of sonification design," *International Symposium on Computer Music Multidisciplinary Research*, pp. 742–757, October 2013. (Cited on pages 91, 96 and 164.)

[196] S. Shelley, M. Alonso, J. Hollowood, M. Pettitt, S. Sharples, D. Hermes, and A. Kohlrausch, "Interactive sonification of curve shape and curvature data," in *Haptic and Audio Interaction Design*, pp. 51–60, Springer, 2009. (Cited on page 91.)

[197] S. C. Peres, "A comparison of sound dimensions for auditory graphs: Pitch is not so perfect," *Journal of the Audio Engineering Society*, vol. 60, no. 7/8, pp. 561–567, 2012. (Cited on page 91.)

[198] R. M. Church and W. Meck, "A concise introduction to scalar timing theory," *Functional and neural mechanisms of interval timing*, pp. 3–22, 2003. (Cited on page 93.)

[199] E. H. Weber, *De Pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae*. CF Koehler, 1834. (Cited on page 93.)

[200] J. Meinilä, P. Kyösti, T. Jämsä, and L. Hentilä, "WINNER II channel models," *Radio Technologies and Concepts for IMT-Advanced*, pp. 39–92, 2009. (Cited on page 94.)

[201] G. Mowbray, J. Gebhard, and C. Byham, "Sensitivity to changes in the interruption rate of white noise," *The Journal of the Acoustical Society of America*, vol. 28, no. 1, pp. 106–110, 1956. (Cited on pages 97, 98 and 164.)

[202] G. G. Wang, "Definition and review of virtual prototyping," *Journal of Computing and Information Science in Engineering (Transactions of the ASME)*, vol. 2, no. 3, pp. 232–236, 2002. (Cited on page 109.)

[203] S. Ullrich and T. Kuhlen, "Haptic palpation for medical simulation in virtual environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 617–625, 2012. (Cited on page 109.)

[204] F. Tendick, M. Downes, T. Goktekin, M. Cenk, M. C. Cavusoglu, D. Feygin, X. Wu, R. Eyal, and M. Hegarty, "A virtual environment testbed for training laparoscopic surgical skills," *Presence*, vol. 9, pp. 236–255, 2000. (Cited on page 109.)

[205] T. Kohler, K. Matzler, and J. Füller, "Avatar-based innovation: Using virtual worlds for real-world innovation," *Technovation*, vol. 29, no. 6, pp. 395–407, 2009. (Cited on page 109.)

[206] M. M. Tseng, J. Jiao, and C.-J. Su, "Virtual prototyping for customized product development," *Integrated Manufacturing Systems*, vol. 9, no. 6, pp. 334–343, 1998. (Cited on page 109.)

[207] OS Community, "Blender". URL: `http://www.blender.org` [accessed 10/18/2014]. (Cited on page 110.)

[208] T. Rick, A. von Kapri, and T. Kuhlen, "A virtual reality system for the simulation and manipulation of wireless communication networks," in *Virtual Reality Conference*, pp. 111–114, IEEE, 2011. (Cited on page 111.)

[209] LIMSI, "Evolutive Virtual Environment". URL: `http://www.limsi.fr/venise/EVEsystem` [accessed 10/18/2014]. (Cited on page 115.)

[210] B. Williams, S. Bailey, G. Narasimham, M. Li, and B. Bodenheimer, "Evaluation of walking in place on a wii balance board to explore a virtual environment," *ACM Transactions on Applied Perception (TAP)*, vol. 8, no. 3, pp. 1–14, 2011. (Cited on page 115.)

[211] B. N. Walker and R. Stanley, "Thresholds of audibility for bone-conduction headsets," in *Proceedings of the International Conference on Auditory Display (ICAD)*, (Ireland), pp. 218–222, 2005. (Cited on page 122.)

[212] W. H. Tuttlebee, *Software defined radio: enabling technologies*. John Wiley & Sons, 2003. (Cited on page 124.)

[213] R. Farrell, M. Sanchez, and G. Corley, "Software-defined radio demonstrators: An example and future trends," *International Journal of Digital Multimedia Broadcasting*, pp. 1–12, 2009. (Cited on page 124.)

[214] J. Mitola III, "Software radios: Survey, critical evaluation and future directions," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 8, no. 4, pp. 25–36, 1993. (Cited on page 124.)

[215] J. Mitola, *Cognitive Radio—An Integrated Agent Architecture for Software Defined Radio*. PhD thesis, Royal Institute of Technology (KTH), 2000. (Cited on page 124.)

[216] J. J. Moskal, *Interfacing a reasoner with heterogeneous self-controlling software*. PhD thesis, Northeastern University Boston, 2011. (Cited on page 124.)

[217] S. Glass, V. Muthukkumarasamy, and M. Portmann, "A software-defined radio receiver for APCO project 25 signals," in *Proceedings of the International Conference on Wireless Communications and Mobile Computing*, pp. 67–72, ACM, 2009. (Cited on page 124.)

[218] B. Bloessl, M. Segata, C. Sommer, and F. Dressler, "An IEEE 802.11 a/g/p OFDM receiver for GNU Radio," in *Proceedings of the second workshop on Software radio implementation forum*, pp. 9–16, ACM, 2013. (Cited on page 124.)

[219] W. H. Tuttlebee, "Software radio technology: a european perspective," *Communications Magazine, IEEE*, vol. 37, no. 2, pp. 118–123, 1999. (Cited on page 124.)

[220] J. Mitola, *Software Radio*. Wiley Online Library, 2003. (Cited on page 124.)

[221] G. Baldini, T. Sturman, A. R. Biswas, R. Leschhorn, G. Godor, and M. Street, "Security aspects in software defined radio and cognitive radio networks: a survey and a way ahead," *Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 355–379, 2012. (Cited on page 124.)

[222] A. Costin and A. Francillon, "Ghost in the air (traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices," *Black Hat USA*, pp. 1–12, July 2012. (Cited on page 124.)

[223] G. Jones, "Mobile menace: why SDR poses such a threat," *Network Security*, no. 6, pp. 5–7, 2012. (Cited on page 124.)

[224] M. Ettus, "USRP user's and developer's guide", Ettus Research LLC. URL: http://http://www.olifantasia.com/gnuradio/usrp/files/usrp_guide.pdf [accessed 10/18/2014]. (Cited on page 124.)

[225] H. Long, "FUNcube, software defined radio dongle receiver". URL: http://www.funcubedongle.com [accessed 10/18/2014]. (Cited on page 124.)

[226] Osmocom, "DVB-T dongle receiver based on RTL-SDR". URL: http://sdr.osmocom.org/trac/wiki/rtl-sdr [accessed 10/18/2014]. (Cited on page 124.)

[227] Mathworks, "USRP2 receiver in Simulink". URL: http://www.mathworks.fr/discovery/sdr/usrp.html [accessed 10/18/2014]. (Cited on page 124.)

[228] National Instruments, "NI USRP: Software Defined Radio Platform". URL: http://www.ni.com/usrp/ [accessed 10/18/2014]. (Cited on page 124.)

[229] N. Kim, N. Kehtarnavaz, and M. Torlak, "LabVIEW-based software-defined radio: 4-QAM modem," in *Proceedings of The World Multi-Conference on Systemics, Cybernetics and Informatics*, vol. 10, pp. 54–61, 2007. (Cited on page 124.)

[230] N. M. Anas, H. Mohamad, and M. Tahir, "Cognitive radio test bed experimentation using USRP and Matlab/Simulink," in *Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, pp. 229–232, IEEE, 2012. (Cited on page 124.)

[231] R. Dhar, G. George, A. Malani, and P. Steenkiste, "Supporting integrated MAC and PHY software development for the USRP SDR," in *1st Workshop on Networking Technologies for Software Defined Radio Networks*, pp. 68–77, IEEE, 2006. (Cited on page 124.)

[232] E. Blossom, "GNU radio: tools for exploring the radio frequency spectrum," *Linux Journal*, pp. 4–8, June 2004. (Cited on page 124.)

[233] D. C. Tucker and G. A. Tagliarini, "Prototyping with GNU Radio and the USRP-where to begin," in *Proceedings of the Southeastcon Conference*, pp. 50–54, IEEE, 2009. (Cited on page 124.)

[234] M. Dickens, B. P. Dunn, and J. N. Laneman, "Design and implementation of a portable software radio," *Communications Magazine, IEEE*, vol. 46, no. 8, pp. 58–66, 2008. (Cited on page 124.)

[235] R. Gandhiraj and K. Soman, "Modern analog and digital communication systems development using GNU Radio with USRP," *Telecommunication Systems*, pp. 1–15, 2013. (Cited on page 124.)

[236] F. Quitin, U. Madhow, M. M. U. Rahman, and R. Mudumbai, "Demonstrating distributed transmit beamforming with software-defined radios," in *International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–3, IEEE, 2012. (Cited on page 124.)

[237] B. Seeber, "Software Defined Radio Direction Finder, Pseudo Doppler application". URL: http://wiki.spench.net/wiki/SDRDF [accessed 10/18/2014]. (Cited on pages 124 and 138.)

[238] J. D. Reed, *Approaches to Multiple-source Localization and Signal Classification*. PhD thesis, Virginia Polytechnic Institute and State University, 2009. (Cited on pages 124 and 138.)

[239] M. Puckette, "Pure data: another integrated computer music environment," *Proceedings of the Second Intercollege Computer Music Concerts*, pp. 37–41, 1996. (Cited on page 125.)

[240] G. Holzmann, "Earplug, binaural filter based on KEMAR impulse measurement for Pd". URL: http://puredata.info/downloads/earplug [accessed 10/18/2014]. (Cited on page 130.)

[241] M. G. SMITH, "Avalanche transceiver recovery skills tips and traps," *GearWorld*, pp. 1–14, 1998. (Cited on page 134.)

[242] C. Oestges, "Radio channel models for search-and-rescue missions into collapsed structures," in *Proceedings of the URSI International Symposium on Electromagnetic Theory (EMTS)*, pp. 1015–1018, IEEE, 2013. (Cited on pages 134 and 163.)

[243] N. Ayuso, J. Cuchí, F. Lera, and J. Villarroel, "A deep insight into avalanche transceivers for optimizing rescue," *Cold Regions Science and Technology*, vol. 102, pp. 32–41, 2014. (Cited on page 134.)

[244] S. Kantesaria, *Software Defined Radar*. PhD thesis, MIT Lincoln Laboratory, 2011. (Cited on page 138.)

[245] A. Akindoyin, M. Willerton, and A. Manikas, "Localization and array shape estimation using software defined radio array testbed," in *Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 189–192, IEEE, June 2014. (Cited on page 138.)

[246] S. P. Smith and D. Trenholme, "Rapid prototyping a virtual fire drill environment using computer game technology," *Fire Safety Journal*, vol. 44, no. 4, pp. 559–569, 2009. (Cited on page 140.)

[247] B. N. Walker and J. Lindsay, "Using virtual environments to prototype auditory navigation displays," *Assistive Technology*, vol. 17, no. 1, pp. 72–81, 2005. (Cited on page 140.)

[248] M. Wolter, C. Armbrüster, J. T. Valvoda, and T. Kuhlen, "High ecological validity and accurate stimulus control in VR-based psychological experiments," in *Proceedings of the Eurographics conference on Virtual Environments*, pp. 25–32, 2007. (Cited on page 140.)

[249] M. A. Schmuckler, "What is ecological validity? A dimensional analysis," *Infancy*, vol. 2, no. 4, pp. 419–436, 2001. (Cited on pages 141 and 153.)

[250] I. Gibson, Z. Gao, and I. Campbell, "A comparative study of virtual prototyping and physical prototyping," *International Journal of Manufacturing Technology and Management*, vol. 6, no. 6, pp. 503–522, 2004. (Cited on page 141.)

[251] J. J. LaViola Jr, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000. (Cited on page 142.)

[252] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence: Teleoperators and virtual environments*, vol. 7, no. 3, pp. 225–240, 1998. (Cited on page 142.)

[253] M. J. Kearns, W. H. Warren, A. P. Duchon, and M. J. Tarr, "Path integration from optic flow and body senses in a homing task," *Perception*, no. 31 (3), pp. 349–374, 2002. (Cited on page 144.)

[254] H. Liu, "Method and system for determining a location of a wireless transmitting device and guiding the search for the same", US Patent, US 20040029558 A1, 2003. (Cited on page 156.)

[255] J.-G. Remy, "System and a method for locating a mobile terminal, in particular for rescuing in distress, and a device for alerting a corresponding mobile terminal", US Patent, US 20040033796 A1, 2003. (Cited on page 156.)

[256] Netline, "Emerency preparedness and emerency communications network solutions". URL: `http://www.netlinetech.com/products/cellular/emergency-preparedness.html` [accessed 01/11/2015]. (Cited on page 156.)

[257] Linkabit, "Ma-308 handheld DF antenna". URL: `http://www2.l-3com.com/linkabit/products/sigint_systems.htm` [accessed 01/11/2015]. (Cited on page 156.)

[258] iSis, "Smartphone based intelligent rescue system". URL: `http://www.isis-application.com` [accessed 01/11/2015]. (Cited on page 156.)

[259] Ranplan, "iBuildNet - Indoor radio network planning tool". URL: http://www.ranplan.co.uk/ [accessed 01/06/2015]. (Cited on page 162.)

[260] A. Schmitz, T. Rick, T. Karolski, T. Kuhlen, and L. Kobbelt, "Simulation of radio wave propagation by Beam Tracing.," in *EGPGV*, pp. 17–24, 2009. (Cited on page 163.)

[261] G. Parseihian, B. F. Katz, and S. Conan, "Sound effect metaphors for near field distance sonification," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 6–13, 2012. (Cited on page 163.)

[262] R. N. Shepard, "Circularity in judgments of relative pitch," *The Journal of the Acoustical Society of America*, vol. 36, no. 12, pp. 2346–2353, 1964. (Cited on page 164.)

[263] D. Poirier-Quinot, D. Touraine, and B. F. Katz, "BlenderCAVE: A flexible open source authoring tool dedicated to multimodal virtual reality," in *5th Joint Virtual Reality Conf (JVRC)*, (Orsay), pp. 19–22, 2013. (Cited on page 167.)

[264] D. Felinto and M. Pan, *Game Development with Blender*. Cengage Learning, 1st ed., 2013. (Cited on page 167.)

[265] T. Lentz, I. Assenmacher, M. Vorländer, and T. Kuhlen, "Precise near-to-head acoustics with binaural synthesis," *Journal of Virtual Reality and Broadcasting*, vol. 3, no. 2, pp. 1860–2037, 2006. (Cited on page 170.)

[266] M. Rébillat, E. Corteel, and B. Katz, "SMART-I2 "Spatial Multi-user Audio-visual Real-Time Interactive Interface"," in *Audio Engineering Society Convention*, pp. 1–16, 2008. (Cited on page 170.)

[267] W. Chen, A. Plancoulaine, N. Férey, D. Touraine, J. Nelson, and P. Bourdot, "6DoF navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms," in *19th ACM Symp on Virtual Reality Soft and Tech (VRST)*, (New York), pp. 111–114, ACM, 2013. (Cited on page 171.)

[268] H. Li, "Création et contrôle de l'avatar dans la plateforme BlenderCAVE," Master's thesis, Université Paris-Sud, 2014. (Cited on page 171.)

[269] W. Chen, N. Ladeveze, C. Clavel, D. Mestre, and P. Bourdot, "User cohabitation in multi-stereoscopic immersive virtual environment for individual navigation tasks," in *IEEE Virtual Reality (IEEE VR)*, p. (accepted), 2015. (Cited on page 171.)