Aparicio, J., Álvarez, F.J., Hernández, A. & Holm, S., 2022, "A survey on acoustic positioning systems for location-based services", IEEE Transactions on Instrumentation and Measurement, vol. 71, art. no. 8505336, pp. 1-37.

Available at http://dx.doi.org/10.1109/TIM.2022.3210943

A Survey on Acoustic Positioning Systems for Location-Based Services

Joaquín Aparicio, *Member, IEEE,* Fernando J. Álvarez, *Senior Member, IEEE,* Álvaro Hernández, *Senior Member, IEEE,* and Sverre Holm, *Senior Member, IEEE*

***Abstract*—Positioning systems have become increasingly popular in the last decade for location-based services such as navigation and asset tracking and management. As opposed to outdoor positioning, where the Global Navigation Satellite System became the standard technology, there is no consensus yet for indoor environments despite of the availability of different technologies, such as radiofrequency, magnetic field, visual light communications, or acoustics. Within these options, acoustics emerged as a promising alternative to obtain high-accuracy low-cost systems. Nevertheless, acoustic signals have to face very demanding propagation conditions, particularly in terms of multipath and Doppler effect. Therefore, even if many acoustic positioning systems have been proposed in the last decades, it remains an active and challenging topic. This paper surveys the developed prototypes and commercial systems that have been presented since they first appeared around the 1980s, to 2022. We classify these systems into different groups depending on the observable they use to calculate the user position, such as the Time-Of-Flight, the Received Signal Strength, or the acoustic spectrum. Furthermore, we summarize the main properties of these systems in terms of accuracy, coverage area and update rate, among others. Finally, we evaluate the limitations of these groups based on the link budget approach, which gives an overview of the system's coverage from parameters such as source and noise level, detection threshold, attenuation, and processing gain.**

***Index Terms*—acoustic positioning systems; link budget; RSS; spread-spectrum; Time-Of-Flight; ultrasounds;**

I. Introduction

THE Global Navigation Satellite System (GNSS) provides nowadays a standard positioning solution for outdoor environments. However, it does not work properly indoors, due to the high attenuation and complex reflections experienced by the transmitted electromagnetic waves in buildings. As people nowadays can spend up to around 90% of their time indoors [1], [2], this situation strengthens the need for a similar solution to that available outdoors. As a consequence, indoor positioning systems and indoor Location-Based Services (LBS) are experiencing a growing interest by both academics and industry: A recent market report predicted a compound annual growth rate of 22.5% between 2020 and 2025 [3]. This growing interest has been observed in a variety of use cases, such as: healthcare, where they increase productivity by locating more efficiently people or assets [4], [5]; retail stores, where they are used to evaluate shopper behavior, marketing techniques and inventory management [6], [7]; provide navigation of people and autonomous vehicles in search and rescue missions [8], [9]; cargo tracking and fleet management in logistic and transportation services [10], [11]; provide wayfinding, multimedia guides and content recommendation for augmented reality applications in museums [12], [13]; smart environments, where it can support ambient assisted living in addition to navigation guidance [14], [15]; and more recently, contact tracing applications in a pandemic, aiding the health services by automatically identifying close contacts [16], [17].

In all these cases, an underlying positioning system is required to calculate the user/target position. Currently, there are different technologies that have been employed to that end, such as Wi-Fi [18]–[20], Bluetooth [21]–[23], Ultra-Wide Band (UWB) [24]–[26], Radio Frequency Identification (RFID) [27]–[29], cell towers [30]–[32], camera [33]–[35], visible light communications (VLC) [36]–[38], Inertial Measurement Units (IMU) [39]–[41], magnetic fields [42]–[44] and acoustics [45]–[47]. All these technologies have advantages and disadvantages, and in general, different considerations about the desired accuracy, deployment cost, infrastructure availability, area to monitor, privacy concerns and the availability of local companies providing that service, will decide for one particular technology.

Systems based on radiofrequency (RF) waves such as Wi-Fi, Bluetooth or cell towers, have a long useful range and can benefit from existing infrastructure, so they can be interesting for LBS in large environments. However, they usually have position errors around a few meters [48], and they can locate the target in a different room or floor, as RF waves propagate through walls. Within RF systems, UWB can provide better accuracies, but their dedicated infrastructure can be costly and be affected by scalability issues [49]. Acoustic systems can also provide high positioning accuracy, as the nominal sound speed in air of 343 m/s is six orders of magnitude smaller than the speed of light, which translates to lower errors when converting time measurements into distances. These systems usually work at frequencies between some kHz and 50 kHz, so the required sampling frequencies are not as high as those required for RF systems, reducing the cost

This research was funded in part by The Research Council of Norway under project 269614, and in part by the Spanish Government and the European Regional Development Fund (ERDF) through project MICROCEBUS (RTI2018-095168-B-C51/C54) and project POM (PID2019-105470RA-C33).

J. Aparicio was with the Department of Informatics, University of Oslo, 0373 Oslo, Norway (e-mail: joaqapar@ifi.uio.no).

F. J. Álvarez is with the Department of Electrical Engineering, Electronics and Automation, University of Extremadura, 06006 Badajoz, Spain (e-mail: fafranco@unex.es).

Á. Hernández is with the Electronics Department, University of Alcalá, 28805 Alcalá de Henares, Spain (e-mail: alvaro.hernandez@uah.es).

S. Holm is with the Department of Physics, University of Oslo, 0371 Oslo, Norway (e-mail: sverre.holm@fys.uio.no).

of the associated hardware to be deployed. Moreover, microphones had been readily available in Commercial Off-The-Shelf (COTS) smartphones for some time, as opposed to the lack of embedded UWB sensors in those devices, so nowadays people do not need to carry any additional dongles. Therefore, acoustic positioning systems are considered to be less costly to deploy than UWB systems. Another advantage of acoustic systems with regard to RF systems is that acoustic waves do not propagate through walls, so it is easy to distinguish between adjacent rooms and floors. Also, they do not suffer from challenging light conditions, as opposed to camera-based and VLC systems, and they have no position drift, like IMU systems [49]. Systems based on the measurement of the Earth magnetic field are sensitive to changes in the operation area, such as adding or removing metal structures [49], which have no effect on acoustic systems. All in all, acoustic positioning systems are very versatile and usually provide better accuracy than RF systems, including easy room and floor identification, so they are particularly relevant when accuracy is the most important factor, as can be precise navigation of unmanned vehicles or unequivocal room presence.

Nevertheless, acoustic positioning systems also have disadvantages:

- The small sound speed means that these systems are more sensitive to the Doppler effect caused by the movement between emitter and receiver.
- Multipath is originated by the nature of wave propagation and the corresponding multiple reflections in a confined environment. Its influence depends on the complexity of the room.
- The high absorption ir air limits the range between emitters and receivers.

The first two effects worsen the performance of acoustic positioning systems, and thus, some of the latest research deal with the mitigation of these phenomena [50]–[52]. The high absorption experienced by acoustic waves implies that more beacons are needed to provide enough coverage in large environments, compared to RF systems.

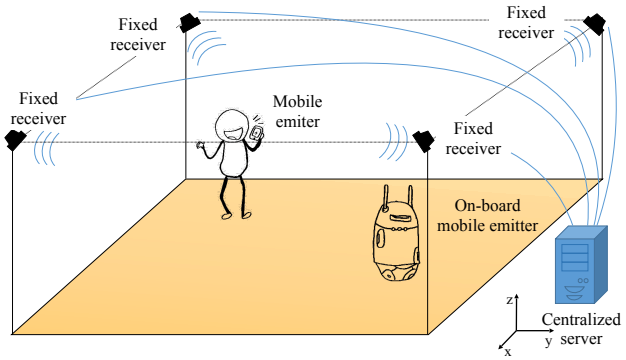
A. Design considerations for acoustic indoor positioning systems

There are different requirements that must be taken into account when designing an acoustic indoor positioning system. The most relevant ones are:

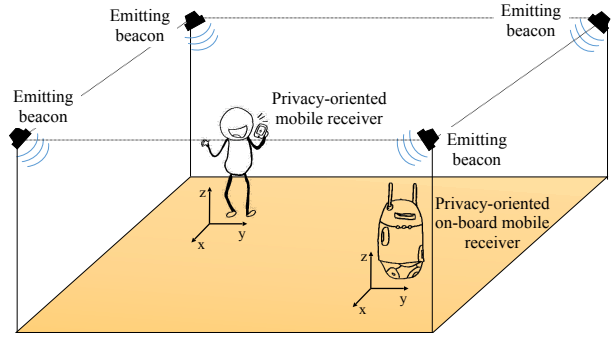
- Synchronization: Some systems require all nodes (emitters and receivers) to be synchronized, whereas others do not. Synchronized systems use additional technologies, such as RF communication links, so all nodes are aware of the transmission times and identity of the transmitters. These systems calculate the Time-Of-Flight (TOF) of the acoustic signal as the difference in the arrival times of the acoustic signal relative to a synchronization pulse, or a known emission time given by the RF link. Non-synchronized systems do not require the target to be synchronized with the other nodes. As these systems cannot calculate TOF, they use one arrival as a reference, and calculate Time-Difference-Of-Arrival (TDOA) instead.
- Accuracy: This is the error in the obtained position with regard to a ground truth. A high accuracy (low error) is obviously desired, although how much is needed is determined by the application. The accuracy depends on the acoustic frequency, and on the distribution of transmitters in the environment and the position of the receiver, an effect that is known as Dilution of Precision (DOP) [53].
- Update rate: It is the number of position updates per second, given in Hz. A high update rate is required for systems that have to track continuously a moving target, but an asset management system can operate with a lower update rate, since the targets will not often move. There is a trade-off with power consumption.
- Coverage area: This is the size of the system operating area, which in indoor environments is commonly the room size where the system is deployed. The actual coverage is affected by the DOP, and therefore, it depends on the number of transmitters/receivers and their distribution over the environment.
- Privacy: Positioning systems can be classified into two groups, as illustrated in Fig. 1. In centralized positioning systems (Fig. 1a), a central unit calculates the position of the user to be located. In privacy-oriented systems (Fig. 1b) the user calculates the position, without communicating it to a central server. The intended application, together with compliance with data privacy laws, will decide: An asset management and tracking system in a warehouse should be centralized, whereas an LBS for people in a department store could be designed as privacy-oriented to safeguard the user's privacy.
- Deployment effort: The desired coverage area and the measurement technique will influence how costly, both economically and in time, it is to deploy and maintain the system.
- Multi-user capability: The total number of users that can be located depends on the nature of the transmitter (the infrastructure or the users), and on additional techniques and communication protocols that can be applied, such as assigning different codes to different targets (Code-Division Multiple Access, CDMA), or assigning them different time slots (Time-Division Multiple Access, TDMA).
- Scalability: Depending on the measurement technique and desired accuracy, it can be more or less costly to expand the system to new areas. For example, a system offering room-level localization based on received signal strength typically needs one beacon per room, so it is easier to extend to more rooms than a more accurate system based on TOF, as they need several beacons per room.

B. Survey methodology

Here we review acoustic positioning systems in the context of LBS for indoor and outdoor environments, developed by both academic institutions and the industry, between the 1980s until 2022. Although there are several reviews of indoor



(a) Centralized positioning system



(b) Privacy-oriented system

Figure 1: Illustration of a centralized (a), and a privacy-oriented (b) localization system.

positioning systems [54]–[71], they cover different technologies, so they do not provide a more in-depth coverage of acoustic systems, compared to this work. Other reviews have focused on acoustic sensing in a broad sense [72], but they have the same issue as acoustic indoor localization is one application among many. The closest review that we have found is [47], which categorized acoustic indoor localization systems into two main groups, absolute and relative range systems. The authors focused on the different measurement techniques rather than on performance, so no information was given about key figures of merit such as coverage area, update rate or accuracy. As opposed to the aforementioned reviews, we focus here on the performance provided by the positioning systems. In order to do so, we consider only those systems in which enough information is given about the design, operation, and evaluation by means of experimental results. We make an exception for commercial systems, as they typically do not provide detailed information about signal design or experimental results. Nevertheless, we believe they should be included in this survey, as they are relevant for the field, so we gather them into their own section. As we focus on LBS applications, this review does not include acoustic systems designed for other applications such as sound source localization [73], [74], Simultaneous Localization and Mapping (SLAM) [75], [76], human-computer interfaces [77]–[79], motion-capture [80], [81] or collision avoidance [82].

We classify the acoustic positioning systems into different groups, based on which observed magnitude, or observable, has been used to calculate the position. The most common observables are TOF, Received Signal Strength (RSS), and the acoustic spectrum, although there are other observables that can be leveraged too. We think that this structure gathers positioning systems that employ similar algorithms, and that exhibit similar advantages and disadvantages. Inside each group, we follow a coarse chronological order and we describe the main features of the systems, the experimental design and the provided results, so the reader can put these results into context. At the end of each group, a table summarizes some important properties, such as update rate and accuracy, when this information is available. Some of these properties, such as the accuracy, have no consensus when reported in the literature. Therefore, we specify the default criteria for each table, and annotations indicate when a different figure of merit is used. By default, area refers to room size, unless specified otherwise. As for privacy, we consider the intended final system and not the prototype status, which in many cases are evaluated on a computer. Additionally, we compare the expected coverage of each group based on the link budget, which relates such performance with magnitudes such as source level, system bandwidth, and noise level.

Section II covers systems based on TOF measurements, whereas RSS systems are reviewed in Section III. Systems that use the acoustic spectrum and other observables are gathered into Section IV, whereas commercial systems are described in Section V. Section VI compares the different groups based on their link budgets, and Section VII gathers the lessons learned and main challenges that remain to be solved. Table I gathers the main abbreviations that have been used in this work.

II. Time of Flight

The measurement of TOF to or from different beacons has been the most exploited technique to calculate the target position in indoor environments. The TOF is the time it takes for the transmitted signal to reach the receiver, and it depends on the distance between the two, and the wave speed. The idea of sending ultrasonic pulses for TOF estimation can be traced back to the 1970s, where spark generators were used in ultrasonic digitizing systems [83]. This idea was later popularized with the introduction of the pulse-echo systems from Polaroid and Yodel Technologies [84], [85]. These systems measured the round-trip TOF, from which the TOF could be obtained just dividing by 2. One of the earliest examples was presented in 1986, and aimed to localize the end-effector of a robot in an industrial environment, up to a range of 3 m [86].

To obtain the TOF, a synchronization between the transmitter and receiver is assumed. This approach requires a minimum of 3 beacons to calculate the 3D position of the target. If no synchronization is available, a positioning system that consists of many beacons can use one of them as a time reference, and calculate instead the TDOA. Unsynchronized systems therefore require one extra beacon to calculate the target position, compared to synchronized systems. Fig. 2 shows a transmission/reception example for a synchronized system, where the transmitted signal is shown in blue, and

Table I: Main abbreviations.

Abbreviation	Meaning
ADC	Analog-to-Digital Conversion
AOA	Angle-Of-Arrival
BLE	Bluetooth Low Energy
BPSK	Binary Phase-Shift Keying
CCK	Complementary Code Keying
CDMA	Code-Division Multiple Access
COTS	Commercial Off-The-Shelf
CSMA	Carrier-Sense Multiple Access
CSS	Complementary Set of Sequences
DOA	Direction-Of-Arrival
DOP	Dilution of Precision
DSSS	Direct-Sequence Spread Spectrum
EKF	Extended Kalman Filter
FDMA	Frequency-Division Multiple Access
FFT	Fast Fourier Transform
FHSS	Frequency-Hopping Spread Spectrum
FPGA	Field-Programmable Gate Array
FSK	Frequency Shift Keying
GCC-PHAT	Generalized Cross-Correlation with Phase Transform
GNSS	Global Navigation Satellite System
ID	Identifier
IEKF	Iterated Extended Kalman Filter
IFW	Interference-Free Window
IMU	Inertial Measurement Unit
LBS	Location-Based Services
LED	Light-Emitting Diode
LFM	Linear Frequency Modulation
LOS	Line-Of-Sight
LPS	Local Positioning System
LS	Loosely Synchronous
LTE	Long-Term Evolution
MFCC	Mel-Frequency Cepstral Coefficients
MSE	Mean Squared Error
MUSIC	Multiple Signal Classification
NLOS	Non-Line-Of-Sight
OFDM	Orthogonal Frequency-Division Multiplexing
OOK	On-Off Keying
PDA	Personal Digital Assistant
PDR	Pedestrian dead-reckoning
PLL	Phase-Locked Loop
QPSK	Quadrature Phase-Shift Keying
RF	Radio-frequency
RFID	Radio-Frequency Identification
RFNN	Region-Feature Neural Network
RH	Relative humidity
RMS	Root Mean Square
RMSE	Root-Mean-Square Error
RSS	Received Signal Strength
RTOF	Round-trip Time-Of-Flight
SLAM	Simultaneous Localization and Mapping
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
SSID	Service Set Identifier
TDMA	Time-Division Multiple Access
TDOA	Time-Difference-Of-Arrival
TOF	Time-Of-Flight
USB	Universal Serial Bus
UWB	Ultra-Wide Band
VLC	Visible Light Communication
WLAN	Wireless Local Area Network

the received signal is depicted in red. The TOF between them is approximately 3.5 ms. The TOF/TDOA is then converted into distance by multiplication with the sound speed, which depends primarily on temperature, and to a lesser extent on humidity, pressure, and frequency, and these variations may need to be compensated for. Appendix A will describe this dependence in more detail, but a nominal value is 343 m/s for a temperature of 20°C. After the calculation of ranges, the target position is commonly obtained by different techniques, such as multilateration or triangulation [87], [88] or signature matching (fingerprinting) [89]. Alternatively, the position equations can be solved by non-linear least squares algorithms [90], [91], or by a Kalman filter [92].

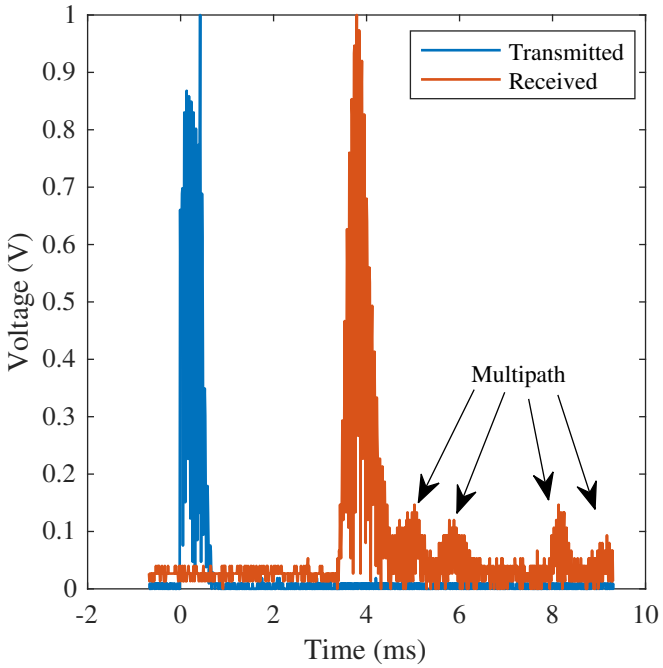


Figure 2: Plot of normalized amplitudes for a transmitted 40 kHz pulse of 16 periods (blue), and the received echo (red), measured with an Agilent DSO6014A oscilloscope with 8-bit resolution.

We have divided the systems based on TOF measurement into four different subgroups, depending on how they use the information, the signal bandwidth compared to the center frequency, and the use of encoding sequences, as illustrated by Fig. 3. The selected categories are depicted in blue blocks. The first distinction has been made between systems based on the received echo structure and the rest, since the former use the full echo pattern, as opposed to the detection of the first arrival employed by the other systems. This latter group has been subdivided attending to the bandwidth of the transmitted acoustic signal into narrowband and wideband systems. In this work, we have considered a system to be narrowband when the bandwidth of the transmitted signal is less than 10% of the carrier frequency, and wideband otherwise. The bandwidth has been calculated for a 3 dB fall from the signal spectrum maximum. Acoustic transmitters are commonly characterized by a small bandwidth, as opposed to RF systems. Narrowband systems typically use a bandwidth of tens or hundreds of hertz, and wideband systems have a bandwidth of few kilohertz at best. This situation impose different transmission strategies in both groups, where narrowband systems usually employ several cycles of sine waves and therefore longer signal durations, whereas wideband systems can transmit more complex signals of shorter duration. This situation has an impact on the influence of multipath on the received signal and the update

rate of the system. Finally, attending to the complexity of the transmitted wideband signals, wideband systems have been further subdivided into uncoded and coded systems, where the latter transmit modulated binary sequences, as opposed to short pulses or chirps.

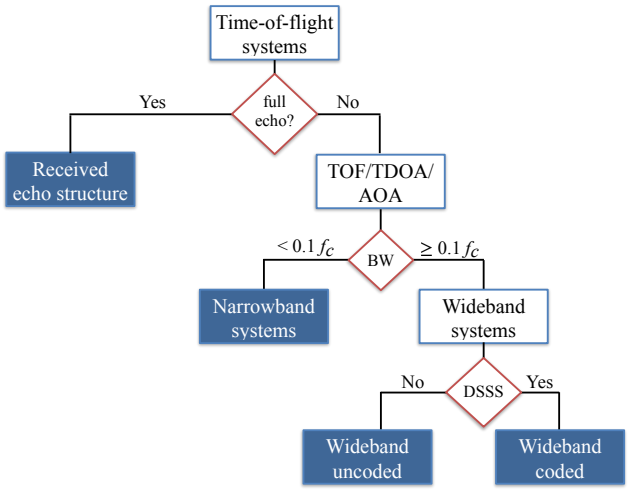


Figure 3: Classification followed in this work of Time-of-Flight systems based on echo information used, bandwidth BW, center frequency f_c, and the use of binary sequences.

A. Received echo structure

Considering the received signal (red line) in Fig. 2, the TOF method uses the timing of either the rising flank, or the signal's maximum amplitude. But as it can be observed in the figure, the received signal contains more information that could be utilized by the system. In indoor environments, this received signal typically contains reflections from the room surfaces, such as walls, ceiling and floor. This reception from different paths is known as multipath, and in principle, it would be possible to implement 3D positioning with only a single transmitter/receiver pair, if this additional information could be leveraged. Each location of interest in the room is assigned a unique signature, and a signature matching process provides the positioning.

The first step is to characterize the room, which can be done in different ways:

1) By obtaining a map of the room with a sonar device.
2) By using an acoustical model of the room and the devices to compute the signatures.
3) By measuring signatures on a fine grid throughout the room.

Sonar devices arose as a popular method at the end of the 1970s and beginning of the 1980s in order to avoid collisions in robot navigation [93]–[95]. As opposed to stereo vision navigation systems, which were computationally expensive, sonar range measurements provided denser information with lower computational cost, although with lower accuracy. Using sonar devices, mobile robots obtained maps of the surroundings in systems developed in the 1980s at the University of Tsukuba, Yale University, Carnegie-Mellon University, General Motors Research Laboratories or more recently, CCS Robotics [96]–[100]. These maps could be later used in SLAM systems.

The first approach of those mentioned above was used by MIT [101]. They mounted a Polaroid ultrasonic rangefinder on a mobile robot to obtain a 360° sonar contour of rooms. After data acquisition, different algorithms were applied to extract line segments (walls) from the sonar data, to pair those segments to walls, to remove erroneous locations that imply sonar rays crossing walls, and finally a heuristic algorithm could choose between location candidates based on the greatest amount of the sonar contour in contact with the walls, based on available maps of the rooms. The system was tested against a total of 24 experiments in three different rooms, achieving a 70.8% success: in 17 tests the location error was less than 30.5 cm, with an orientation error less than 5°, whereas in the other 7 tests either the location was wrong, or no location was provided by the algorithm.

Reflective cylinders at known locations were used as artificial landmarks in [102] to calculate the position and orientation of a mobile robot. The robot was equipped with 18 Polaroid transducers, which were used to sense the environment. From the TOF measurements, different regions of constant distance were obtained, corresponding to reflections from the environment, such as walls, but also from the cylinders. All these regions provided candidate locations for the cylinders. Since the cylinders' locations were known, the theoretical distance between them was compared with the one obtained from the potential candidates, and the candidate pair providing the minimum error was used to calculate the position and the orientation of the robot. Experimental tests following different trajectories provided location errors below 3 cm, and orientation errors below 2°.

Learning algorithms were used as early as 1996 to identify rooms based on sonar measurements from mobile robots [103]. In this experiment, a K2A base from Cybermotion, and a Nomad 200 mobile robot platform were used to collect data by using rings of ultrasonic sensors distributed in two different heights and locations around the robots body. The echoes from the sonar were translated to pixels in a 2D map, and a Region-Feature Neural Network (RFNN) was used to classify those measurements to the available maps of the rooms. Ten different rooms were considered, where the robots would first collect the data, and then train the network. Processing time could be reduced by using the features previously learned by the other robot platform, obtaining then classification rates between 77 and 88%.

More recently, [104] provided room-level localization using a sonar device on a ActivMedia Pioneer 3 AT mobile robot platform. The ring-projection histogram method was used as feature descriptor since it provides rotation invariance but it is sensitive to scale. The key idea was to leverage the room-level sparsity and to combine multiple sets of readings in the model. Experimental tests were conducted in 10 different areas. Samples were collected during 10 days for training, and test samples were collected afterwards during three days. Under situations where the rooms stayed in similar condition as during the training phase, accuracy was around 90%. This accuracy reduced to approximately 75% under significant

changes in room conditions.

The aforementioned second signature matching method was tested in [105]. In this work, a single base station transmitter and a mobile receiver were considered. Different 3D candidate positions were then evaluated by computing and comparing their expected TOF signatures with the measured one at the mobile receiver. According to Fig. 2, the TOF signatures would include in this case the arrivals at 5, 6 and 8 ms, in addition to the first one at 3.5 ms. A key part of this approach was to calculate the Line-Of-Sight (LOS) range, which was provided by the first TOF and a synchronization process via RF link. Several signal processing steps were also involved, such as attenuation compensation. The results were obtained with a base station placed near the ceiling, and taking 20 measurements at different positions in the room with the receiver at a height of 1.3 m. Most of the times, the error was below 20 cm.

The RoomSense positioning system was developed as an app for Android phones [106]. As the transmitter and receiver were the same device, it can be considered an infrastructure-free system. It calculated the impulse response of rooms at different locations by means of m-sequences with a duration of 0.68 s, thus following the third approach. The Mel-frequency cepstral coefficients (MFCC) where extracted then from the impulse response, and training data were classified accordingly. A Samsung Galaxy SII was used in the experiments, where 20 rooms in two different buildings were considered, under controlled conditions (windows and doors closed, and no furniture moved). A Support Vector Machine classifier provided the estimated position in less than 1 s, using the leave-one-sample-out method. Room classification rate was 98.2%, which dropped to 85.1% if the orientation of the smartphone was not trained. Similarly, when considering fewer locations inside a room for the training, performance dropped to 49.8%. For sub-room-level accuracy, the classification rate was 96.4%, which dropped to 51.3% when the orientation was not trained. When increasing the noise level so the Signal-to-Noise Ratio (SNR) dropped from 50 to 10 dB, the classification rate decreased to 66.6 and 65.9% for room-level, and sub-room-level accuracy, respectively.

Table II summarizes the main characteristics and results of these systems. Note that the system's accuracy refers to the absolute error when a numeric value is given, and that a dash (–) means that the information was not available.

B. Narrowband systems

A narrowband spectrum is obtained when transmitting a pulse consisting of several cycles of a sinusoidal wave. One of the main advantages of narrowband systems is the simplicity of the detector, which is usually designed as a more or less complex version of a tone or envelope detector, such as the one depicted in Fig. 4. The price to pay for this simplicity is trifold, though: low accuracy, poor update frequency, and high sensitivity to in-band noise. Alternatively, linear frequency modulated signals, or chirps, could be designed with a narrow bandwidth of less than 10% of the carrier frequency, so those systems fall into our narrowband category. In those cases though, the detector usually performs a correlation process, such as the one depicted in Fig. 5.

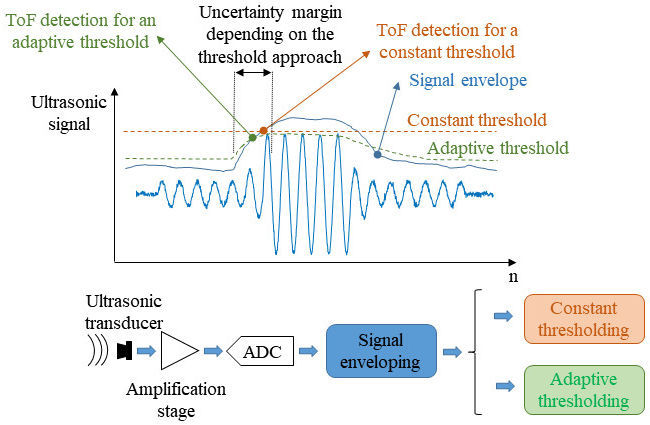


Figure 4: Typical detector of a narrowband system.

MELODI was one of the first positioning systems [107]. It had three beacons fixed at known locations, and a receiver device that consisted of three different pairs of sensors, which measured the Angles-of-Arrival (AOA) to the beacons. Frequency-Division Multiple Access (FDMA) was considered in transmission to avoid collisions, with each beacon transmitting at a different frequency, whereas each pair of sensors was matched to only one beacon. By obtaining the phase difference at each pair of sensors, the three angles to the transmitters could be obtained and a triangulation operation provided the location of a robot carrying the receiver device, as well as the heading. In experimental tests placing the robot in different locations inside a room, absolute location errors were usually below 5-10 cm under good coverage conditions.

The system by Monash University also consisted of different beacons fixed at known positions, which transmitted 2.5 ms ultrasonic pulses [108]. An array of eight receivers was used to detect those pulses. The beacons were controlled by a central system and transmitted sequentially to avoid collisions, where one of them, which transmitted two pulses instead of one, was used as a reference at the receiver. This way, TDOA could be used in the array to calculate the position and the orientation. An early prototype provided location errors around 5 cm, and orientation errors around 10°, using six beacons deployed in a room. A later version [109] employed an Iterated Extended Kalman Filter (IEKF) for the positioning of a robot. It used the dead-reckoning on board the robot to get the change of angular velocity and speed between two consecutive measurements, so drift effects could be reduced. Once the IEKF had converged, multipath arrivals could be also filtered out by observing the observation error in the filter equations. Maximum errors of 10 cm where obtained in a trajectory followed by a mobile robot equipped with the receiver array.

Cricket used an RF signal to synchronize the emission and to identify the transmitting beacon [110]. Concurrently with the RF signal, an ultrasonic pulse was emitted by the beacon, which was detected by a simple tone detector circuit at the receiver. This receiver was synchronized to the beacon via the RF signal, being then able to calculate the TOF.

Table II: Summary of systems based on received echo structure.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
M.I.T.	1987	55	No	–	30.5	–	Yes	[101]
Gold Star Industrial Systems KAIST	1995	50	No	–	3	31	Yes	[102]
North Carolina State Univ.	1996	–	–	–	Room-level	256	Yes	[103]
Eindhoven Univ. of Tech. Philips Research Lab.	2003	40	Yes	–	20-100	≈ 29	Yes	[105]
RoomSense	2013	Audible	No	–	Sub-room-level	15-65	Yes	[106]
Tsinghua Univ.	2017	–	No	–	Room-level	20.1-311	Yes	[104]

Only coarse-grained location information to the nearest beacon was obtained. In a later work [111], the authors evaluated a hybrid configuration of the system, where the target became active when detecting high uncertainty in its position tracking, provided by the state of an Extended Kalman Filter (EKF). In this hybrid approach the target pinged the beacons with an RF message and an ultrasonic pulse, and the beacons replied then also via RF with their measured distance to the target. This procedure allowed to reset the target's EKF. Under this hybrid configuration, 90% of the errors were below 20 cm when the target moved at a speed of 0.78 m/s, and below 47 cm when moving at 1.43 m/s. In both cases the hybrid approach reduced the errors obtained using a purely passive system, especially at low speeds.

A similar approach was followed by the positioning system from the University of Bristol [112], where four ceiling mounted beacons were synchronized to a receiver by an RF signal. These beacons transmitted pulses sequentially, and they were detected by the receiver, which was attached to a hand-held device. The three shortest TOF were used to calculate the user location, and the application interface averaged the last two positions, in order to smooth the trajectory shown on screen. The maximum error at the center of the test room was 10 cm, degrading up to 25 cm when placing the receiver at the perimeter. A later version of this prototype system removed the RF signal [113], reducing the receiver size to a single ultrasonic microphone and its associated amplification circuit. Knowledge of the periodic transmissions from the beacons and their locations was leveraged in an EKF to calculate the user location. The obtained standard deviation in a 3-minutes test with a moving user was between 5 and 10 cm.

WALRUS consisted of two components: a server beacon, and a mobile client [114]. The server beacon was a computer with attached speakers, which could send packets over a Wi-Fi access point containing information about the room, while transmitting a 10 ms sound pulse through the speakers. The mobile client had an integrated microphone and should be able to receive Wi-Fi packets. Every time the receiver detected a Wi-Fi location packet, it recorded an audio window from the microphone and looked for the pulse. A localization algorithm matched then the reception times of the Wi-Fi packet with the ultrasound detection, considering the potential reception of multiple packets from different rooms. For 2 server beacons correct detection was achieved almost 100% of the time, but this number dropped to 84% for 6 server beacons, and approximately 45% for 25 server beacons.

Beep was designed to locate COTS mobile devices such as Personal Digital Assistants (PDA) and cellphones based on the transmission of short pulses [115]. As opposed to the previous approaches, it transmitted in the audible frequencies with an intensity similar to a cellphone ring in order to make use of the best sensitivity band of the receivers. These receivers were connected to a desktop computer, and synchronized with the user either by a Wireless Local Area Network (WLANBeep), 3G (3GBeep), or a pure acoustic protocol (PureBeep). Upon a user transmission, the receivers measured the TOF, transformed them to distances, and reported them to a server, which then applied a lateration algorithm using the three smallest distances, sending the calculated 2D position back to the user. In the experimental tests, 100 measurements were taken at 48 equally distributed points throughout a room. An iPAQ HP 5550 was used with the WLANBeep and PureBeep protocols, and an Audiovox SMT 5600 cellphone was employed for 3GBeep. The results were similar for all protocols. Under quiet conditions and considering 90% of the measurements, WLANBeep and PureBeep provided positioning errors below 70 cm.

The low-cost privacy-oriented synchronous system presented by the Bristol Robotics Laboratory [116] was based on an architecture of eight beacons that periodically emitted ultrasonic pulses using Murata transducers. This system was aimed to track and perform loop control of drones in 3D, and it used an RF signal to synchronize the beacons with the receiver, that detected the ultrasonic emission by means of a high-gain amplifier plus a microcontroller. In the experiments, the beacons were placed on the floor, and were used to track a drone. For fixed positions of the drone, a standard deviation around 0.3 cm was obtained for the (x, y) coordinates, whereas a standard deviation of 2.3 cm was obtained for the z coordinate, for measurement periods of 500 s.

FLASH provided the 2D location of sensors based on the reception of audible pulses generated in turns by buzzers [117]. These pulses were received by the sensors to be located, which detected the TOF on the peak-to-peak amplitude of the signal, combined with a moving-average filter and a dynamic threshold. Experiments were conducted using two nodes as

beacons, obtaining the 2D position of another 10 nodes with a maximum error of 36 cm.

SpiderBat combined distance with AOA measurements, obtained from the TOF [118]. This positioning system was designed as an extension board to be attached to the nodes of sensor networks, in order to provide location capabilities. All nodes were synchronized with an RF link, and they sent the collected range and AOA information from neighbouring nodes to a central computer for the calculation of the locations. This calculation was done in two steps. In the first one, the nodes with unknown position got information about the closest anchor node, and the measured range and AOA provided a rough location estimation. A second step using the ranges in a least mean square algorithm provided the final position. Starting with only one anchor node to locate three more nodes in 2D, standard deviations of the calculated locations varied between 2.2 and 5.7 cm in an indoor environment, and between 37.7 and 51.6 cm outdoors.

The system proposed by the Autonomous University of Madrid used semi-custom electronics based on Xilinx Spartan 3/3A Field-Programmable Gate Arrays (FPGA) in the design of the emitter and receiver modules of a synchronous system [119]. A three-beacon structure was used to sequentially emit squared pulses, which were captured, amplified and digitalized at the receiver. The main novelty was the use of a state machine to accomplish the detection of the ultrasonic emissions in order to measure the TOF. This solution avoided the delay and uncertainty typically introduced by envelope detection. The authors reported average errors around 3.3 cm in the estimation of 2D coordinates, and an average error of 5.6 cm for the height.

The prototype presented by the National Institute of Informatics and Hokkaido University used a combination of two pulses at different frequencies to distinguish between the different beacons used in transmission [120]. As three beacons were used to calculate the TDOA, a total of six pulses were used with a separation of 500 Hz. These pulses were received by a smartphone, and the TDOA were extracted by evaluating the phases of the different components. Two points were used to evaluate the prototype, where 100 measurements were taken for each of them. The average error was around 10 cm. Increasing the noise level slightly worsened the accuracy to around 12 cm. A later prototype, called SyncSync, employed Light-Emitting Diode (LED) signals for synchronization [121]. Experiments were conducted in a room equipped with a LED floodlight of 56 white LED, while turning the ceiling lights off. Four speakers transmitted the acoustic signals following an FDMA scheme, and the errors were calculated at 10 different points, measuring 150 positions at each of them. The mean error ranged between 4.7 and 17 cm, with a maximum standard deviation of 1 cm.

A coordinator node was connected to a computer in the TELIAMADE system, in order to control a set of wireless battery-powered end nodes equipped with ultrasound transducers and ZigBee chips, which provided synchronization and communication capabilities [122]. The transducers transmitted pulses of 1 ms using a TDMA protocol to avoid collisions. After propagation, the mobile node received these pulses and calculated the TOF after amplification, filtering, sampling and quadrature detection. The mobile node sent this information back to the coordinator, which fed it to the computer in order to calculate the position of the node. In the experiments, four end nodes were deployed in the center of a room to avoid multipath from the walls, and the mobile node was placed at 17 different locations in two different heights (10.8 and 62.8 cm). Three hundred measurements were taken at each point, obtaining locations with a Root-Mean-Square Error (RMSE) of 0.93 cm and 0.98 cm for the points at 10.8 cm and 62.8 cm height, respectively.

ASSIST was designed to locate smartphone users in a centralized fashion [123]. It ran an Android app, which connected to an evaluation unit using the Long-Term Evolution (LTE) network. It then obtained a unique identifier (ID), and transmitted a 50 ms chirp with a bandwidth of 1 kHz. This signal was received by a set of fixed microphones, which applied a high-pass filter, amplification and digitalization, calculated the TDOA and extracted the user ID. TDOA and user ID were later sent to the evaluation unit, which calculated the position. In an experiment using 7 receivers placed at a height of 1 m, and a person walking along a predefined track, the obtained average deviation from the track was 34 cm. A later improvement used the inertial sensors embedded in the smartphone by means of a Kalman filter [124]. This addition allowed to obtain a location when the acoustic signal was lost, and to correct the drift when there was an acoustic position fix. Two experiments were conducted in a hangar, considering six receivers and different Non-Line-Of-Sight (NLOS) conditions, obtaining RMSE values for the locations between 16.9 and 22 cm, when using the inertial sensors and a Fuzzy Inference System.

The so-called ALPS system had the same purpose, but it followed a privacy-oriented approach [125]. It consisted of three or more beacons, which contained an ultrasound transceiver board and a Bluetooth Low Energy (BLE) daughter board. The beacons employed a TDMA protocol with slots of 100 ms to transmit 50 ms chirps between 20 and 21.5 kHz for ranging, where up and down chirps were transmitted alternatively to further reduce the interference terms between consecutive transmissions. The beacons also transmitted periodic BLE packets so the user could synchronize to the TDMA cycle. Another novelty was the addition of a room mapping algorithm assisted by the user, in order to generate the floor plan by using landmarks such as corners, locating the beacons in the floor plan afterwards after taking a few measurements. An iPhone 5S was used in the experiments, and the data were processed offline. An EKF fused the range data with compass and step values from the phone, obtaining location errors less than 30 cm for 90% of the measurements under LOS conditions.

One of the most recent narrowband systems is from Harbin Institute of Technology [126]. The main novelty of this work was a TOF detection algorithm based on the least squares estimation of the maximum signal amplitude. The system consisted of different anchor nodes deployed at known positions acting as receivers and connected to a server, and a mobile node as a transmitter. Upon detection of an RF synchronization signal, the anchor nodes calculated the TOF and sent this

information to the server for position calculation. The system was tested by a robot following small circular trajectories at low speed to avoid multipath and Doppler, achieving a maximum error of approximately 1 cm.

Table III summarizes the main results and configuration parameters from the narrowband positioning systems. Note that accuracy refers here to the maximum error considering 90% of the measurements in the reported experiments. A dash (–) means that the information was not available.

C. Wideband, uncoded systems

As has already been mentioned, positioning techniques based on lateration require the determination of the TOF from each emitter to the receiver. Therefore, the more accurate the determination of TOF is, the more accurate the final position is. For that purpose, wideband transmissions have been applied to improve the accuracy in TOF estimation, as signal bandwidth is inversely proportional to the standard deviation of the TOF estimator due to receiver noise [127].

Typically, there are two ways to obtain a wideband spectrum: 1) by transmitting very short pulses, since bandwidth is inversely proportional to pulse duration; and 2) by transmitting long signals, spanning a wide bandwidth. The first option would make use of a detector such as the one depicted in Fig. 4, so it does not fully make use of the benefits of wideband signals, as it can be sensitive to in-band noise. The improvement in accuracy is commonly achieved by the second option. In this case, transmissions usually consist of Linear Frequency Modulated (LFM) signals, or chirps, and the receiver employs a correlator, or matched-filter, to look for the transmitted signals, as illustrated in Fig. 5. These filter provides a peak when the transmitted signal is detected, compressing its energy in a narrow interval around its arrival time, increasing then the range resolution and the robustness to noise. This pulse compression technique for the determination of echoes appeared in ultrasonic and sonar systems coming from radar applications [128], and in telecommunications, it is also known as Direct-Sequence Spread Spectrum (DSSS) when used to demodulate spread spectrum signals.

In summary, compared to narrowband systems, wideband uncoded systems provide the following main benefits:

- Higher accuracy in the determination of TOF.
- Higher immunity against noise, being possible to deal with significantly reduced SNR.

Short pulses to obtain wideband transmissions were used in pioneer works such as the ultrasonic location system from the University of Cambridge and Olivetti and Oracle Research Laboratory (ORL) [129]. A target emitted a 50 μs pulse that was detected at fixed receivers. The system operation was controlled by a central computer that sent a reset signal to the receivers via a serial network, and synchronized emitter and receivers via RF. The system architecture was formed by 16 receivers uniformly distributed in a square grid, and errors below 10 cm for 90% of the measurements were reported.

The Spanish National Research Council (*Consejo Superior de Investigaciones Científicas*, CSIC) also presented an early prototype of a wideband uncoded positioning system [130].

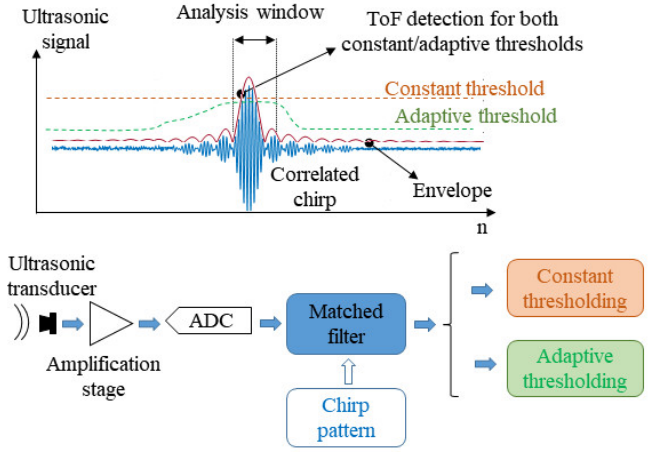


Figure 5: Example of a detector for a wideband uncoded LPS employing a chirp.

The novelty of this work was the use of a spark generator mounted on a mobile robot, which was used as a transmitter. At certain time intervals, the spark generated both acoustic and electromagnetic pulses, where the maximum amplitude in the acoustic spectrum was around 40 kHz, although there was energy leaking into the audible bands. Different elements were connected to the same acquisition board in the receiver: 1) an RF antenna detected the electromagnetic wave and provided synchronization of the system; 2) three ultrasound receivers to calculate the TOF, which were mounted on a wall; and 3) a temperature sensor placed close to the receivers was used to correct for the potential variation in the speed of sound. The received acoustic waves were stored in a computer and processed offline. The calculated position deviated from the given trajectory by a maximum of 3 cm.

A later work from CSIC focused on outdoor environments, and it was used to locate archaeological findings [131]. This was, to the best of the authors' knowledge, the first system specifically designed for outdoor operation that included a wind compensation algorithm. There were three targets in this system: two installed in a portable rod, and one placed at a fixed position to compensate for the wind effect. The targets emitted several cycles of a 25 μs pulse, smoothly modulated in amplitude. This signal was detected by correlation with a reference signal captured from a typical reception, and the target position was estimated by minimizing a cost function using the Levenberg-Marquardt recursive algorithm. Accuracies around 0.4 cm were obtained when considering eight beacons installed in the top and central positions of four fixed posts, with no wind conditions. This accuracy worsened to 1 cm for maximum wind speeds of 5 m/s.

The ZUPS system was design to be deployed in extended environments, such as a residence for elderly people, in order to monitor their behavior and ring an alarm when needed [132]. The design consisted of a set of beacons deployed at known fixed positions that transmitted chirps following a TDMA protocol, and of mobile tags carried by the users. All nodes were synchronized by ZigBee, forming a sensor

Table III: Summary of narrowband systems.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
MELODI	1983	25-40	No	–	5-10[1]	4-6[†]	Yes	[107]
Monash Univ.	1989	40	No	0.25-0.33	5[1]	10	Yes	[108]
Cricket	2000	40	Yes	≈ 4	Sub-room level[2]	–	Yes	[110]
Cricket v2	2004	–	No	–	$\approx$ 20-47	$\approx 4^{\dagger}$	Yes	[111]
Univ. of Bristol	2001	40	Yes	–	10-25[1]	$\approx$ 27	Yes	[112]
Univ. of Bristol v2	2003	40	No	5	5-10[3]	2[†]	Yes	[113]
WALRUS	2005	21	Yes	–	Room-level	–	Yes	[114]
WLANBeep	2006	4.01	Yes	–	70	207	No	[115]
Bristol Robotics Lab.	2007	40	Yes	4-20	3-4[1]	36-84	Yes	[116]
FLASH	2009	2.3	Yes	0.1	11-36[1]	25	Yes	[117]
SpiderBat	2011	40	Yes	–	2.2-51.6[3]	50-60	No	[118]
Autonomous Univ. of Madrid	2012	40	Yes	1.66	3.3-7.3[4]	$\approx 2^{\dagger}$	Yes	[119]
National Inst. of Informatics Hokkaido Univ.	2013	14.75-17.25	No	–	10-12[4]	$0.04^{\dagger}$	Yes	[120]
SyncSync	2015	10.75-15.25	Yes	–	4.7-17[3]	$0.52^{\dagger}$	Yes	[121]
TELIAMADE	2013	40	Yes	1.25	≈ 1[5]	$\approx 0.6^{\dagger}$	No	[122]
ASSIST	2015	18-20	No	–	25-34[4]	80	No	[123]
ASSIST v2	2016	–	–	3.3	8.9-18[4]	$\approx$ 128	No	[124]
ALPS	2015	20-21.5	Optional	–	30-50	$\approx$ 72	Yes	[125]
Harbin Inst. of Tech.	2017	40	Yes	5	≈ 1[1]	≈ 0.2-$0.8^{\dagger}$	No	[126]

[1] Maximum error.
[2] Granularity of approximately 1.5 m^2.
[3] Standard deviation.
[4] Average error.
[5] RMSE.
[†] Area of the grid/trajectory.

network where a coordinator node was responsible for starting the localization process, and then to gather all the information from the nodes. This system also included a NLOS detection algorithm to improve the measurements used in the localization process, which reduced the effect of blocked signals on the obtained accuracy [133]. The system was evaluated in a test area with 6 beacons, and the location of a tag was calculated at 750 different points, obtaining a 2D error below 5 cm, and 3D error below 10 cm, for 99% of the measurements.

Wideband uncoded systems have also been applied to locate smartphone users. One example, applied to the particular context of driving safety, was presented by Stevens University of Technology and Rutgers University, where the objective was to detect if a phone was located in the driver or passenger side inside a car [134]. To that end, the phone connected to the internal Bluetooth network and sent the pulses to be transmitted by the car stereo speakers. These speakers employed a TDMA scheme to avoid collisions, and calculated in which side of the car the phone was by measuring the relative delay obtained between the left and right speaker signals. Two phones (iPhone 3G and Android Developer Phone 2) were tested inside a non-moving car, and the Android phone was also tested on a highway. The correct classification of driver or passenger side, considering a calibration procedure that took into account the dimensions and layout of the speakers inside the cars, varied between 92 and 98%, using a two-channel stereo system and placing the phone in the front seats.

The Whistle system consisted of a set of receivers, which listened to a chirp in the audible band transmitted by the target to be located [135]. The novelty of this work was the lack of synchronization in the receivers to calculate the TDOA. In order to obtain the TDOA they used two consecutive measurements: one by the target, and the other by one of the receivers (base node). This procedure removed the need of a common clock, and it was used together with an outlier removal algorithm and redundant receivers. Several

experimental tests were conducted both outdoors and indoors, using a smartphone as the transmitter, and seven others as receivers, including one acting as base node. All of them were connected to an access point, which gathered the TDOA and sent this information to a laptop for location calculation. The transmitter was placed in a grid of 16 points, both in 3D and 2D, and 10 measurements were taken at each of them. For the indoor experiment and the 3D grid, 12.5% of the measurements were discarded as outliers, obtaining a maximum error of 28.2 cm for 90% of the measurements. Fewer outliers (0.6%) but slightly higher errors were obtained for the 2D grid, with 90% of the measurements below 29.9 cm.

The LOSNUS positioning system dealt with NLOS measurements by means of a proximity based grouping algorithm, leveraging information from the DOP and TDOA uncertainties [136]. The system consisted of 10 transmitters mounted on opposite walls of a room, which transmitted chirps sequentially together with a transmitter ID. The receiver applied a matched filter to obtain the TDOA and minimize the impact of NLOS outliers with the aforementioned algorithm. Experimental tests were conducted placing the receiver in different points, where some of the direct paths to the transmitters were blocked by an object, or by directly facing the receiver to one of the walls. Positioning errors were usually below 3 cm.

The positioning system for smartphones developed by the University of Aveiro consisted of a synchronized wireless sensor network of acoustic beacons, which transmitted chirps following a TDMA protocol [137]. After detecting the chirps from the beacons by a correlation process, the smartphone used the TDOA to obtain its location in 2D as an optimization problem. The location error was below 13 cm for 95% of the measurements using 4 beacons.

The prototype developed by the University of Perugia in collaboration with the Società Delle Fucine S.R.L. also transmitted chirps in a TDMA scheme [138]. This system was designed with the aim to be deployed at large industrial environments, so their main novelty was to use a set of fixed beacons with known positions, together with a portable beacon grid. This grid could locate itself by performing measurements to the fixed beacons. This configuration would allow to improve LOS conditions with regard to the mobile object to be located. The performance of the beacon grid was evaluated experimentally in a lab. Seven transmitters were used to obtain the position of a mobile node located at four different points. The maximum positioning error was 3.9 mm.

In the context of file-sharing applications on smartphones, the AMIL system was designed to obtain the relative localization of Android devices referred to a master user [139], in order to allow file sharing based on proximity. This master user transmitted a linear chirp that was received by a set of smartphone users. Both the master and users calculated the intervals between successive transmissions and receptions, respectively, in order to calculate TDOA. The key idea in this work was to use virtual anchors by moving the transmitter to different points, forming the shape of a triangle with a base of 61 cm, obtaining then three anchor nodes. The movement and location of the transmitter was tracked by the internal IMU of the phone. By drawing three different triangles, average errors were kept below 50 cm for all the 12 users considered.

ARABIS consisted of a location server and several anchor nodes [140]. The anchor nodes, deployed at known locations, contained an acoustic board attached to a Raspberry Pi 3, and transmitted periodically an acoustic message by chirp modulation, which included the anchor ID and a sequence number. The user detected the chirps and decoded the message, obtained the timestamps and communicated this information back to the location server via network (such as Wi-Fi), which then computed the position based on an iterative Gauss-Newton algorithm. Experimental tests were conducted using a Samsung Galaxy phone, obtaining location errors below 35 cm when using 4 anchor nodes, and below 13 cm when using 8 nodes, both considering 90% of the measurements. A later evolution of the system, called AALTS [141], introduced an asynchronous transmission scheme, the use of orthogonal chirps to increase the transmitted data rate of the acoustic messages, and multiple tones in the preamble to compensate for Doppler shift. 2D localization was evaluated by experimental tests in different rooms and noise conditions, where 4 anchors were used to calculate the location of both static and mobile targets. 90% of the errors were below 22 cm for static targets. This value increased to approximately 60 cm under impulsive loud noises. For a mobile target, 90% of the errors were below 22 cm when following a straight trajectory at low walking speed (0.3 m/s), and below 49 cm for a U-shaped trajectory performed at 1 m/s.

SAILoc used a combination of range and Direction-Of-Arrival (DOA) to calculate the position of a smartphone, by deploying a single array node [142]. This array, whose location was known, consisted of a speaker and multiple microphones forming a linear array. The smartphone transmitted a 50 ms chirp, which was recorded at the array as well as in the smartphone. The array proceeded then to transmit the same chirp back to the smartphone, and the arrival time was also recorded by both sensors. A server gathered all the information, and calculated the range and angle between both nodes. Experiments were conducted in a room and a corridor, at 6 different points recording 50 measurements at each of them. 90% of the errors were below 50 cm in the room, and below 60 cm in the corridor.

The prototype developed by the Mediterranea University of Reggio Calabria consisted of a master unit and a central unit equipped with the beacons [143]. The master unit, connected to a computer, had RF communication capabilities. The central unit stored a 2.56 ms chirp signal, which was fed sequentially to four beacons. The mobile units, synchronized to the master and central units, acquired the received chirps and a correlation process provided the TOF to the beacons. Experimental tests were done by taking 100 measurements at 18 different points. Reported errors were below 0.45 cm for the most favorable point, and below 1 cm for the worst point, considering 90% of the measurements.

The University of Science and Technology of China presented a system consisting of four beacons as transmitters, and a smartphone acting as receiver [144]. In this system the beacons followed a TDMA-FDMA transmission protocol, where two of them emitted simultaneously at a time, avoiding

collisions by using chirps at different frequency bands. Two steps were needed then for a full transmission cycle. Upon receiving the chirps, the smartphone calculated the TDOA, and calculated an estimated position using the shrinking-circle method. A Doppler shift correction based on the knowledge of the previous position was applied to that initial estimate to obtain the final position. Experiments were conducted under static and dynamic conditions. Forty test points were used for the static tests, obtaining absolute errors below 20 cm for 90% of the experiments. Three different trajectories were considered for the dynamic tests, with walking speed between 0.4 and 0.9 m/s. Mean Squared Errors (MSE) varied between 12 and 27 cm.

PALS combined acoustics with pedestrian dead-reckoning (PDR) [145]. The acoustic positions were obtained with a ranging method that iteratively removed multipath components [146]. These acoustic positions were fused in an EKF with data calculated from the inertial sensors embedded in the phone, such as the stride length. Experimental tests were performed using six synchronized beacons and a Huawei Mate 9 as receiver, where the user followed two different trajectories inside the area of interest. 90% of the measurements had errors below 27 cm.

In the system developed by the University of Tsukuba and the National Institute of Technology (Japan) [147], chirps were transmitted by four beacons at known locations using an FDMA scheme to avoid collisions. The system calculated the AOA measured at two pairs of microphones. For the experiments, the receiver pairs were placed facing upwards in a mobile robot, forming a cross. The microphones recorded the received signals, which were later processed offline. The location was calculated by a particle filter, using a likelihood function based on envelope detection of the cross-correlation functions obtained from each of the two microphone pairs. This particle filter also incorporated information from the wheel rotation to estimate the initial position of the particles, as well as velocity and angular velocity. In the harshest experiment with reflections from a wall next to one of the beacons, 90% of the calculated positions showed an error below 23 cm.

Chirps were also used by the system developed by Wuhan University [148]. In this system, synchronized beacons were used to send chirp signals at scheduled times and with different rates and frequency bands, to minimize interference. These chirps were received by a smartphone, and a detection method based on the GCC-PHAT and a time-frequency analysis selected the first arrival path from each transmitter, reducing the effect of reverberation. The estimated TDOA were fed to a least-squares algorithm, which was combined with an EKF for tests under dynamic conditions. Experiments were performed in two different testing areas (hall and parking lot), and using different phones as receivers. Four beacons were used in the hall, and 8 in the parking lot, due to its larger size. For the static tests, 41 points were considered across the hall, whereas 56 points were considered in the parking lot, where both testing areas included NLOS and signal blockage conditions. The results varied notably with the used smartphone: 95% of the positioning errors were below 0.89 m using a OnePlus 6, and below 1.92 using a Pixel 3. Errors increased under dynamic conditions, obtaining maximum errors between 1.6 and 1.85 m, depending on the smartphone, considering 95% of the measurements.

The main novelty of one of the most recent wideband uncoded systems was to leverage the reflected signal from surfaces in order to reduce the number of transmitters down to a single speaker [149]. This speaker transmitted chirps that were received at a smartphone, held sideways to utilize the top and bottom microphones for yaw and pitch estimation. Different outlier detection methods were applied for peak detection and pair selection at the microphones, and arrival angles were used to classify the paths. Experiments were performed in a corridor placing the smartphones on 12 positions, at two different heights. By considering the reflections from the left wall, the right wall and the floor, a single speaker provided 4 peaks that allowed to position the smartphone. The reported errors were below 5.5 cm and 11.4 cm for 90% of the measurements, for a smartphone height of 1.4 and 1.1 m, respectively.

Table IV summarizes the main parameters and the results for the wideband uncoded positioning systems. Note that the accuracy refers to the maximum error considering 90% of the measurements. A dash (–) means that the information was not available.

D. Wideband, coded systems

An alternative technique to obtain a wideband spectrum is to transmit a distinctive sequence of data. At the receiver, a filter matched to the original sequence provides the peak when it is detected.

At the beginning of the 1990s some started to apply the pulse compression technique based on pseudo-random noise to the determination of TOF in ultrasonic sensory systems. Most of them were based on the use of Barker codes to encode the ultrasonic transmissions [150]–[152]. The encoding of ultrasonic transmissions by these sequences always requires the use of a wide bandwidth [153]. This can be a significant issue to deal with, when most commercial ultrasonic transducers are characterized by a bandwidth of few kHz. This bandwidth constraint often results in a reduction of performance, since the correlation functions are not that close to the ideal ones. In order to deal with this problem, a modulation is often proposed, so the final transmitted signal can be adapted in a better way to the frequency response of the sensory system. In previous work, the Binary Phase-Shift Keying modulation (BPSK) has been considered as one of the most suitable options for local positioning systems [154].

The selection of appropriate sequences according to the application is an important point in the design of ultrasonic ranging systems. Whereas Barker codes were a popular first choice, subsequent studies employed different kinds of sequences [155], [156]. Kasami codes seem to provide better auto- and cross-correlation functions (see Fig. 6) [157], with lower sidelobe-to-mainlobe ratio (the lower this value, the better the reception and discrimination of transmissions). Nevertheless, it is also necessary to keep in mind the computational load required by the corresponding matched filters in reception.

Table IV: Summary of wideband uncoded systems.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
ORL	1997	40	Yes	5	5-10	75‡	No	[129]
CSIC	1999	40	Yes	40	3[1]	≈ 0.4†	No	[130]
CSIC-Outdoor	2005	40	Yes	–	0.4-1[1]	16†	No	[131]
ZUPS	2008	–	Yes	1.66	5-10[2]	80	No	[132]
Stevens Univ. of Tech. Rutgers Univ.	2011	16-18, 18-20	No	0.5-1	92-98%[3]	3.2-3.8	Yes	[134]
Whistle	2011	2-6	No	–	21.9-38.1	81	No	[135]
LOSNUS	2014	35-65	No	–	5[4]	36	Yes	[136]
Univ. of Aveiro	2015	18-22	No	2.86	13-16[5]	30†	No	[137]
Univ. of Perugia Società Delle Fucine S.R.L.	2015	38-42	Yes	0.2	0.39	≈15	No	[138]
AMIL	2016	18-22	No	–	≈ 61[1]	15	No	[139]
ARABIS	2017	17.5-21.5	No	0.06	13-35	82.8-225	No	[140]
AALTS	2020	17-22	No	5	25-49	≈ 27.5†	No	[141]
SAILoc	2017	1-3	No	–	50-60	≈ 11-13†	No	[142]
Mediterranea Univ. of Reggio Calabria	2019	30-50	Yes	2	≈ 0.4-1	16	Yes	[143]
Univ. of Science and Technology of China	2019	15-22	No	1	20	≈ 62	Yes	[144]
PALS	2019	–	Yes	–	27	–	Yes	[145]
Univ. of Tsukuba Nat. Inst. of Technology	2019	12-21.5	No	5	19-23	16†	Yes	[147]
Wuhan Univ.	2021	15-21	No	1-2	0.89-1.92[5]	180-900	Yes	[148]
Hokkaido Univ. Nat. Inst. of Informatics	2021	12-18	No	2	5.5-11.4	4	Yes	[149]

[1] Maximum error.
[2] 99% confidence interval.
[3] Percentage of correct detections.
[4] Errors were grouped into categories. A few errors were categorized as > 5 cm.
[5] 95% of measurements.
† Area of the grid/trajectory.
‡ Operating volume (m^3).

From the hardware implementation point of view, sequences derived from Complementary Set of Sequences (CSS) often have efficient structures for correlation, where the number of multiplications and additions can be significantly reduced [158]–[161]. Some work have studied the performance figures of the more common sequences [162].

Other coding schemes such as Loosely Synchronous (LS) codes, derived from CSS, could be remarked [161], [163]. These codes provide perfect auto- and cross-correlation functions inside a certain window around the main lobe (time of arrival), which is called the Interference-Free Window (IFW) -see Fig. 7-. Outside this IFW the correlation values are quite high, potentially degrading the system performance, but these codes are very useful when it is possible to know that all signals should be received in a certain interval inside the IFW. More recently, complex sequences with Doppler resilience such as Zadoff-Chu have been tested for indoor localization [164]. In that work, more advanced modulation techniques as Quadrature Phase-Shift Keying (QPSK) or Orthogonal Frequency-Division Multiplexing (OFDM) were used to transmit Zadoff-Chu sequences with a carrier frequency of 40 kHz, and a bandwidth of 10 kHz. Promising results were obtained when using QPSK-modulated Zadoff-Chu sequences of 257 symbols.

Summing it up, pulse compression techniques based on encoded signals provide some advantages to sensory systems. As wideband uncoded systems, it also provides higher accuracy in TOF determination and higher immunity against noise, compared to narrowband systems. And additionally, wideband coded systems allow for simultaneous emissions with low crosstalk, known as Code-Division Multiple Access (CDMA), which is more difficult to achieve in wideband uncoded systems. Because of these properties, encoding techniques have been applied not only in ultrasonic systems, but also

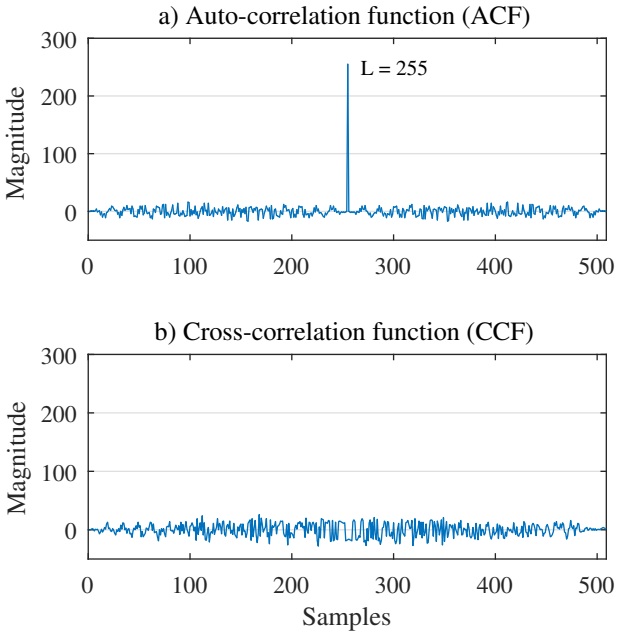


Figure 6: Example of a) auto-correlation function of a 255-bit Kasami code, and b) cross-correlation function of two orthogonal 255-bit Kasami codes.

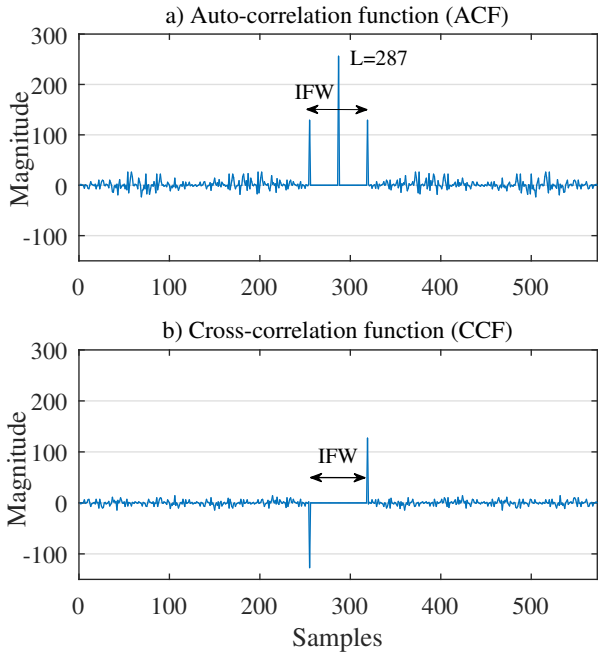


Figure 7: Example of a) auto-correlation function and b) cross-correlation function of a 287-bit LS code.

in other sensory technologies [165]–[169]. Nevertheless, there are also some drawbacks:

- High computational load, even when efficient correlation architectures are available.
- All the analog blocks related to the emitter and the receiver become more complex, since they have to meet certain requirements in linearity and bandwidth.

Taking into account the previous considerations, a general block diagram of the processing usually involved in pulse compression techniques is proposed in Fig. 8 for the emission, and in Fig. 9 for the reception. For the emission module, it is necessary to generate the sequence for each beacon, to modulate it to meet the bandwidth requirements of the sensory system, to convert it into the analog domain, and finally, to amplify it before driving the ultrasonic emitter (beacon). Fig. 8 shows a 255-bit Kasami sequence $c_i[n]$ in a system with four ultrasonic beacons ($i = 1, \ldots, 4$), where the emitted Kasami $e_i[n]$ has a carrier frequency of 40 kHz.

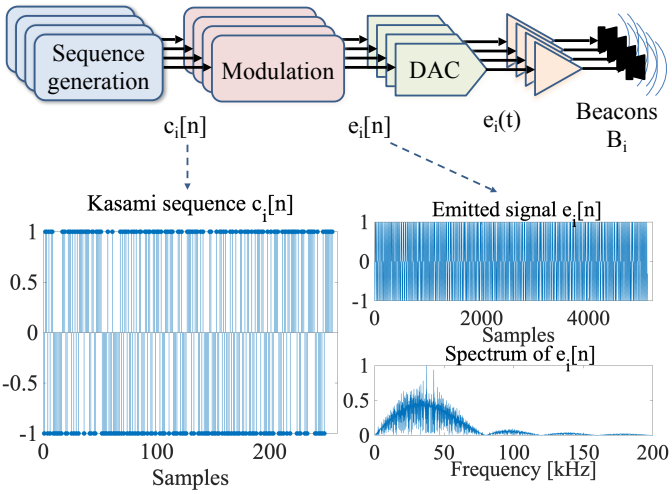


Figure 8: General block diagram for emission in an ultrasonic system based on pulse compression.

In the reception stage, after Analog-to-Digital Conversion (ADC), the second block computes the demodulation. Assuming a BPSK modulation, the computation is described by (1):

$$d[n] = \sum_{k=0}^{M-1} r[k+n] \cdot s_c[k], \tag{1}$$

where $d[n]$ is the demodulated signal, $r[n]$ is the received signal, $s_c[n]$ is the modulation symbol formed by a period of the carrier, and M is the oversampling value, equal to the ratio between the sampling frequency f_S and the carrier frequency f_c: $M = f_S/f_c$.

After demodulation, a correlation is computed to search for the emitted sequences. This process is described in (2), although it can be optimized depending on the kind of sequences applied:

$$t_i[n] = \sum_{k=0}^{L-1} d[M \cdot k + n] \cdot c_i[k], \tag{2}$$

where $t_i[n]$ is the correlation signal for the sequence c_i emitted by the beacon i; $d[n]$ is the demodulated signal; $c_i[n]$ is the searched sequence; L is the length of the sequence $c_i[n]$; and M is the oversampling parameter again.

One of the earliest wideband coded systems was Dolphin [170]. The transmitted signal was encoded with a 511-bit

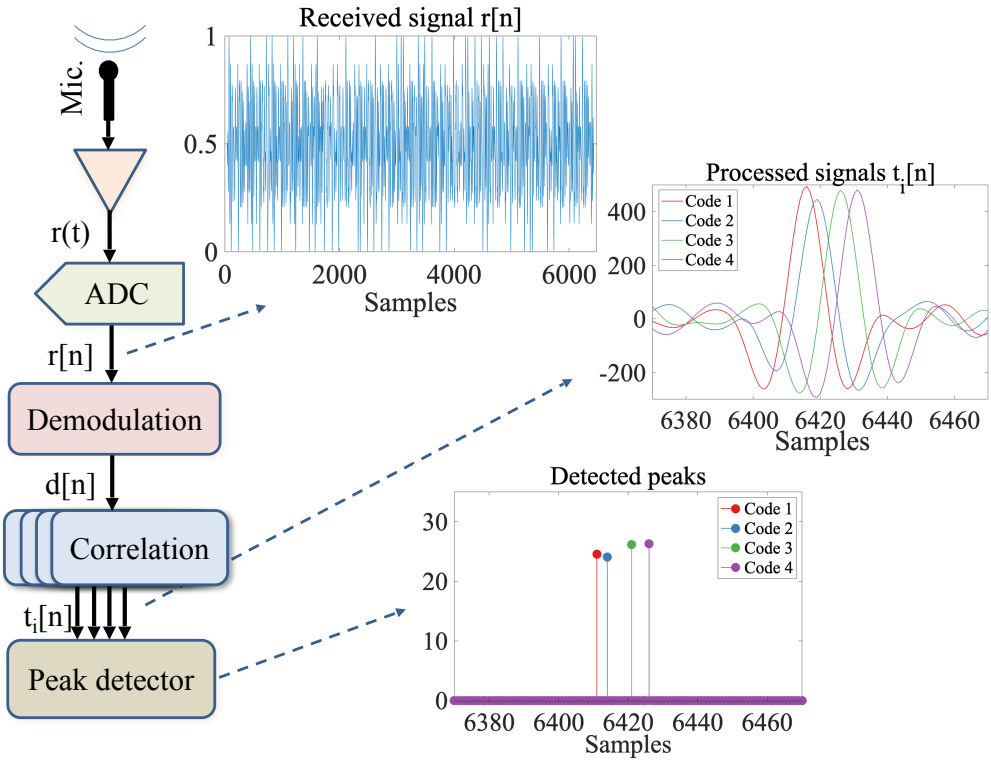


Figure 9: General block diagram for reception in an ultrasonic system based on pulse compression.

Gold code, and a set of eight receivers were deployed in the ceiling. The receivers captured the encoded signal and fed it to a data acquisition card, which performed analog-to-digital conversion, and sent it to a workstation for further processing and position calculation. This setup allowed to achieve errors below 2.34 cm in 3D positioning for 95% of the measurements. In a later work, a privacy-oriented version of Dolphin was presented [171]. In this case, seven units transmitted simultaneously 511-bit Gold codes, which were captured by a mobile receiver. The system was tested in 64 different points, taking 500 measurements at each of them. A Successive Interference Cancellation algorithm was applied to reduce the errors caused by cross-correlation, increasing the number of valid readings. 95% of the 3D position errors were below 4.9 cm when emitters and receiver were synchronized, and below 26.6 cm when they were not.

A similar approach was followed in [172], where a set of 4 beacons with known locations were used to locate a mobile target. In this case, the beacons transmitted a 63-bit m-sequence. Synchronization was achieved by a trigger signal, which allowed to compute to TOF on a computer. An experimental test provided a positioning accuracy of 5 mm.

In contrast to this conventional deployment of several beacons in the room ceiling, the system from Eindhoven University and Philips Research Laboratories consisted of a compact single unit which contained an array of three elements, with a baseline of 11.5 cm [173]. Each element of the array transmitted a unique m-sequence of 127 bits, which was used to phase-modulate a sinusoidal signal. Both the array and the mobile receiver were synchronized by a trigger signal. Experimental results, considering different orientations of the mobile receiver to test for occlusions and the effect of multipath, provided 3D position errors that varied between 1.14 m in the best case, to 3.20 m in the worst one, considering 95% of the measurements.

The 3D-LOCUS system [174] can be considered an extension of the wideband uncoded system for outdoor environments presented in [131]. In this later work, seven beacons were used to locate a receiver, and the system could compensate for temperature gradients and airflows in the room by using one additional node. The transmission consisted of 32-bit Golay codes modulated in BPSK. Twenty-two test points were considered, and more than 100 measurements were taken at each of them. The performance was tested for different emission protocols (CDMA and TDMA), as well as privacy-oriented, centralized and bidirectional configurations. Position errors for 90% of the measures were below 0.41 cm for a privacy-oriented TDMA system, and below 1.37 cm for a centralized CDMA with an airflow present.

University College Dublin developed a location and orientation estimator [175], [176]. In this system, the beacons transmitted Kasami codes by using Frequency-Hopping Spread Spectrum (FHSS), instead of a phase modulation. One beacon was defined by two Kasami codes: one to build the sequence, and the other one was used to generate the hopping pattern. A mobile device was equipped with a circular array of receivers, and it was synchronized with the beacons by means of an RF link. The orientation estimator was based on the calculation of the angle of arrival on the array by applying the Multiple Signal Classification (MUSIC) algorithm. The achieved location accuracy was lower than 1 cm for 95% of the measurements, with maximum errors of 4.5°, 3° and 3.5°

in yaw, pitch and roll, respectively.

Another system that obtained 2D position and orientation was developed by KAIST [177]. It consisted of a group of transmitters at fixed locations, and an array of three receivers that should be mounted on top of a robot. All transmitters and receivers were synchronized by an RF link. The robot selected a transmitter and activated it, and the transmitter replied with its ID, position and temperature via RF, together with an ultrasonic modulated code. This code was detected at the receivers with a matched filter, obtaining the TOF and calculating the 2D coordinates and the heading angle. A later version of the system was tested for the dynamic positioning of a Pioneer P3-DX mobile robot [178]. In this case, the ultrasonic system was fused with the odometric information in an extended Kalman filter. Real tests were conducted using 6 and 10 transmitters, with the robot following different trajectories. When the robot was moving at a speed of 1 m/s, the maximum error was 25.7 cm. In general, the RMSE values for different trajectories and speeds varied between 5.4 and 8.6 cm.

An alternative approach to the typical signal design can be found in the system developed by Carnegie Mellon University [179]. Each beacon was assigned a different ID, which was encoded as a sequence of two Hamming codes. The codes were then modulated by concatenating two upchirps with different rates, in the frequency band between 19 and 24 kHz. These chirps were transmitted after a downchirp that indicated the beginning of the transmission. The experimental setup was based on commercial smart devices (iPhone 3GS), which recorded the received signal and sent it back to a server for processing. At the receiver, a first matched filter detected the downchirps and the TDOA, and knowing the downchirp position, different matched filters were applied subsequently to search for the upchirps, which allowed to recover the transmitter ID. Experimental tests were done with four beacons in two different environments, office and atrium, obtaining errors below 10 and 100 cm, respectively, considering 95% of the measurements.

FX Palo Alto Laboratory and the University of California Irvine exploited the idea of using low-power pseudo-random signals emitted by speakers to provide localization inside rooms [180]. These pseudo-random signals spread throughout the spectrum from 0 to 22.05 kHz, but their low SNR was intended to disturb as less as possible. For the evaluation of the system, six speakers were deployed in the ceiling of a room, and 20 ground truth points were defined inside. A Google Nexus S smartphone running an app was used as the receiver, and sent a request each time it required a position. It recorded the received signal three times at each point, and uploaded the data to a server using the Wi-Fi. There, the TOF were calculated by means of the generalized cross-correlation with phase transform (GCC-PHAT). For pseudo-random signals of 10 s long, the system provided a 3D position with an error below 1.5 m for approximately 85% of the measurements, when using the four best speakers. A privacy-oriented continuous tracker mode was also tested by connecting a microphone to a laptop. Performing a similar study, the system provided 2D positions with errors below 20 cm for approximately 95% of the measurements.

511-bit Gold codes were used in the positioning system tested at the Izmir Institute of Technology [181]. This system consisted of a computer working as a central unit, an external sound card, amplifier, four tweeters and a digital thermometer to accurately calculate the speed of sound. The main novelty of this work was the use of two receivers synchronized with the transmissions: one was placed at the center of the working area, and it was used to calculate transmission delays between the theoretical and the actual transmission time, in order to correct time lags; the other microphone was the target to localize. After cross-correlation, a trilateration algorithm was applied to calculate the position by averaging triplets of distances. The receivers were placed in different points inside one quadrant of the projected area of the tweeters, reporting accuracies below 1.5 cm for 90% of the measurements.

The localization system by Kyoto University, System Watt and NIRO was designed to be used by an agricultural robot in an outdoor setting [182]. A robot was equipped with a transmitter that sent a 1023-bit m-sequence modulated in BPSK, together with a pulse for Doppler compensation and a trigger signal. After reception of the trigger signal, the Doppler shift was calculated and a new replica of the transmitted sequence was generated at the receivers to recalculate the correlation in order to remove the Doppler effect. Experimental tests were conducted by deploying four receivers, and making the transmitter follow a circular trajectory of 1 m radius. After outlier rejection, RMSE location errors were below 2 cm. A more recent, privacy-oriented version was developed by Kyoto University and National Fisheries University [183]. This system was used to locate a quadcopter in a greenhouse, and employed four transmitters at the corners of the greenhouse, and one receiver at the quadcopter, where all nodes were synchronized with an RF link. The best performance was obtained when the transmitters used a TDMA protocol, when requested by the quadcopter. The transmitted signals were 2047 m-sequences modulated in BPSK, which were correlated to obtain the TOF, followed by position calculation by a least squared algorithm. The 2D RMSE errors after 50 repetitions at 8 points inside the greenhouse were 1.6 and 2.3 cm, depending on the noise conditions caused by the thrusters.

Guoguo consisted of beacons deployed in an office environment, connected via Zigbee with a server [184]. In this system, the beacons transmitted a combination of different m-sequences at a center frequency of 18 kHz with 5 kHz bandwidth, so a part of the signal fell into the audible spectrum. A power control system was used to limit the energy of the transmitted signals, so they could not be distinguish from background noise. A smartphone performed a matched-filter operation, calculated the TDOA and sent this information back to the server, which estimated the location. An interesting feature of this system was the use of advanced error pruning techniques based on Root Mean Square (RMS) delay spread or kurtosis analysis to detect outliers or NLOS channel conditions, assigning lower weights to those measurements during the location estimation. An iPhone 4S was used for the experimental tests, obtaining errors that were below 5 and 11 cm for 90% of the measurements, which were taken in 12

different environments.

The system developed by the University of Alcalá was designed for real-time operations by means of FPGA devices, both at the beacons (Xilinx XC2S50E) and receiver (Xilinx XC2S200E) [185]. The beacons were synchronized by a wire, and they transmitted orthogonal 255-bit Kasami codes simultaneously and periodically. A static receiver was placed at four different points in an area of approximately 1 m^2, obtaining location errors in the order of millimeters. This system was later extended to cover an area of 442 m^2 by combining six individual Local Positioning Systems (LPS) deployed at different locations [186]. The receiver, aboard a Pioneer DX5000 robot, used the odometric information when there was no coverage from an LPS, and an H-infinite filter fused both location information when the robot detected the 1023-bit Kasami codes transmitted by the beacons. Inside the coverage area of the LPS, the mean error on the absolute position was below 3 cm. The last version of this positioning system was named LOCATE-US, and consisted of a set of five transducers connected to a central unit deployed in the ceiling [187]. The architecture of the central unit is still based on an FPGA device, consisting of an ARM processor and advanced peripherals to manage the ultrasonic emissions, which can be adapted to any type of modulation, as well as any sequence used to identify the different beacons. A combination of TDMA and CDMA (T-CDMA) was used, transmitting 255-bit Kasami sequences modulated in BPSK in different slots. The system was tested by measuring 100 positions at different points in a grid, obtaining location errors below 20 cm for 90% of the measurements. A recent experiment considered five sets deployed over a large area [188]. Ultrasonic measurements were used together with an external bluetooth-linked inertial sensor, in order to obtain positions in those regions with no coverage from the ultrasonic units. Both sensors were fused in an EKF. A 57-meter-long trajectory was repeated four times at normal walking speed, obtaining maximum errors between 35 and 65 cm, approximately, considering 90% of the measurements.

In [189] the objective was to take advantage of existing infrastructure, so the system could be used through common speakers available in public spaces, employing smartphones as receivers. In the experimental tests, four computer speakers were used as transmitters, and a condenser microphone as a receiver. Different correlation methods, medium access protocols and positioning algorithms were evaluated. The best results were obtained for the transmission of pseudo-random signals, correlated with a GCC-PHAT filter, and fed to a Gauss-Newton algorithm. 95% of the errors obtained in a 23-point grid were below 10 cm.

In the positioning system from the University of Extremadura, the transmissions from four beacons were encoded with 63-bit Kasami sequences, modulated in BPSK and transmitted simultaneously [190]. The receiver was a third generation iPad. The main novelty of this work was the use of a parallel interference cancellation algorithm to deal with the inter-symbol interference caused by the limited bandwidth of the receivers, as well as the multiple-access interference originated from stronger emissions from nearby beacons, compared to far-away ones, which is known as the near-far effect. After detecting the four beacons by correlation, hyperbolic multilateration was used to obtain the location of the receiver. Experimental tests provided errors below 10 cm for SNR above 0 dB. In a more recent work in collaboration with the University of Oslo, the authors exploited a different approach in the signal design to deal with the Doppler effect, near-far effect and low coverage areas [191]. The signals were transmitted sequentially from five beacons, and consisted of three parts: a chirp, a pulse and a distinct 63-bit Kasami code modulated in OFDM. The chirp provided arrival candidates in a window frame containing the signals from all the beacons, the pulse was used to correct the Doppler effect caused by a moving receiver on the received signal, and the correlation with the demodulated Kasami code allowed to validate (or discard) the arrival, and to identify the transmitting beacon. A multipath detection algorithm based on data from previous successful positioning instances identified multipath conditions and corrected erroneous peaks, and the system was also robust to the loss of one or two beacons, by searching through the best combination of the remaining beacons to provide a valid 2D position. Experimental tests were conducted under static and dynamic conditions, using 5 beacons with fixed positions that transmitted sequentially following a T-CDMA protocol. For the static tests, a 49-point grid was considered, and 100 measurements were taken at each of them, obtaining mean errors between 0.2 and 8.9 cm, depending on the test point. Different trajectories were taken next with a mobile robot, and the proposed system was compared with a classic design of up-down chirps. For a 2-meter-side trajectory, the classic system was able to obtain a valid position in 50.19% of the points, whereas the proposed system increased the coverage to 97.89%.

The University of Stuttgart recently presented an opto-acoustic system to be used in an industrial setting for the accurate tracking of tools [192]. A printed circuit board with four transmitters and an infrared LED was attached to the object to be tracked. This object synchronized with a set of fixed receivers by an infrared signal, which was transmitted at the same time than the ultrasonic signal. The ultrasonic signals consisted of 127-bit Gold codes modulated in BPSK, which provided orthogonality to detect the different transmitters in the same board, as well as with other 30 objects. The receivers contained a bank of correlators with different replicas at several carrier frequencies, in order to compensate the amplitude decay of the correlation peak caused by the Doppler effect. Distance measurements were combined with accelerometer and gyroscope data from the object, which allowed to obtain its pose by using a particle filter. When using four transmitters and four receivers in an electrical axis moving at different speeds, the median position error was below 2.6 cm, and the median orientation error was below 10°.

The main parameters that characterize the previous systems are summarized in Table V, together with the obtained results. Note that the system's accuracy refers to the maximum error obtained for 90% of the measurements in the reported experiments. A dash (–) means that the information was not available.

Table V: Summary of wideband coded systems.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
Dolphin	2002	40-60	Yes	–	$\approx$ 1.75-1.9	9.1	No	[170]
Dolphin (privacy-oriented)	2006	40-60	Optional	40	4.9-26.6[1]	9.1	Yes	[171]
Univ. of Magdeburg	2005	39.2-41.2	Yes	–	0.5[2]	6.25†	No	[172]
Eindhoven Univ. Philips Research Lab.	2004	40	Yes	0.3	114-320[1]	28.7	Yes	[173]
3D LOCUS	2009	15	Yes	2-10	0.41-1.37	4†	Optional	[174]
Univ. College Dublin	2009	40	Yes	–	$\approx$ 0.75	8	Yes	[175]
KAIST	2013	35-45	Yes	5	5.4-8.6[3]	38.4-84	Yes	[178]
Carnegie Mellon Univ.	2012	19-23	No	–	10-200	25-400	Yes	[179]
FX Palo Alto Lab. Univ. of California at Irvine	2012	0-22.05	No	2	20	31.6	Optional	[180]
Izmir Inst. of Tech. Joseph Fourier Univ.	2012	7	Yes	–	1.5	3	No	[181]
Kyoto Univ., NIRO System Watt Co., Ltd.	2013	24	Yes	4	≈ 2[3]	≈ 3†	No	[182]
Kyoto Univ. National Fisheries Univ.	2020	24	Yes	1	1.6-2.3[3]	72	Yes	[183]
Guoguo	2016	15-20	No	$\approx$ 1.4	5-15	–	No	[184]
Univ. of Alcalá	2018	34-47	No	10	20	12.6†	Yes	[187]
INESC TEC Univ. of Porto	2016	Audible	Yes	–	2.5-7.5	42	Yes	[189]
Univ. of Extremadura	2018	16	No	2	$\approx$ 10	1†	Yes	[190]
Univ. of Extremadura Univ. of Oslo	2021	36-44	No	5	0.2-8.9[4]	5.76†	Yes	[191]
Univ. of Stuttgart	2020	40	Yes	15	2.6-3.6[5]	0.6†	No	[192]

[1] 95% of measurements.
[2] Absolute error.
[3] Root Mean Square Error (RMSE).
[4] Mean error.
[5] Median error.
† Area of the grid/trajectory.

III. Received signal strength

Received Signal Strength (RSS) systems evaluate the received energy in order to estimate the location of the user, where this calculation can be done with different levels of complexity:

1) The simplest systems can check for the presence or absence of an ultrasonic signal, in order to provide coarse room-level localization.
2) The received energy could be also compared to an existing template measured at different locations (fingerprinting [89]), in order to pick the location with a value closer to the one measured.
3) Alternatively, an attenuation model can be used to estimate the distances between emitters and receivers, followed by a lateration algorithm [87], [88].

These systems can also be interpreted as having coded pulses, which are used for a communication link. This is a key part that is required by all RSS systems, so that each tag or room can be differentiated by transmitting a unique signature. However, unlike the wideband coded systems presented in Section II-D, which integrated the energy over the entire pulse length to obtain a correlation peak and calculate TOF from it, these systems integrate each bit (chip) independently. That is, these systems do not obtain correlation peaks to calculate TOF, only the communication link is needed to identify the bits from the room/tag signature. They are more robust than TOF-wideband systems and cover larger areas, since they use a much narrower bandwidth, having larger tolerance to noise.

Some work on ultrasound communications links have been reported before based on Frequency Shift Keying (FSK) [193]. FSK was compared with other modulation methods such as On-Off Keying (OOK) and BPSK in [194], and the same group later analyzed QPSK [195]. However, in the evaluation of BPSK and QPSK they only considered stationary point-to-point links with continuous communications. For positioning purposes, it is necessary to have multiple transmitters and receivers in the same room, and there is almost always

relative movement. Although other modulation methods may be more efficient for sustained communications, based on the experience from underwater communications when only short burst of data without training sequences are to be sent, FSK has been found to be robust and reliable [196], [197].

As part of the detection process, Doppler shift needs to be estimated. Doppler shifts can be particularly challenging for a narrowband system. This problem is very similar to that experienced by FSK communications in underwater acoustics, where Doppler shifts are also relatively high [196], [197]. Notwithstanding, potentially useful information about the movement can be extracted from the Doppler shift as a side-effect. If $\pm v$ is the maximum velocity, the maximum Doppler shift can be calculated from (3):

$$f_D = \pm f_0 v/c, \tag{3}$$

where v is the velocity component along the ultrasound beam. Considering $v = 6$ km/h or 1.67 m/s, which corresponds to fast-paced walking, and a sound speed $c = 340$ m/s, a Doppler shift of ± 200 Hz is obtained. To solve this issue, the system in [198] derived different frequencies from a 8 MHz master clock. Therefore, dividing this clock frequency by factors 195, 197,..., 205, resulted in frequencies of approximately 41.026, 40.609,..., and 39.024 kHz, i.e. about 400 Hz apart. It was then possible to separate zeroes from ones even when there was maximal Doppler shift. For a 3-bit system, three frequency pairs were used for signaling, one for each bit, where each frequency represented either a value of 0 or 1.

The Doppler bandwidth must be divided in order to make the most from a small detection bandwidth. Bins of size 35 Hz were considered in [198], resulting in approximately 2·400 Hz/35 Hz = 23 bins that had to be checked using a processor based on the Fast Fourier Transform (FFT). A detection in the upper half of the bins was classified as a 1, whereas a detection in the lower half was classified as a 0. Additionally, the bin with the maximum signal provided an estimate of the Doppler shift.

Reusing frequencies might be necessary if more than 3 bits are required for data transmission. Because of the typically low bandwidth of the transducers (usually 10% relative bandwidth), there is a finite number of available frequencies. This solution would work as long as the reuse interval is longer than the reverberation time of the room. A system of this kind was described in [199], where four frequency pairs were used, together with a symbol length of 40 ms and a reuse interval of 160 ms (reuse after four symbols). Even if the data rate of this system was only 25 bits/s, this was enough for many indoor location applications.

An example of the measured velocity component along the ultrasound beam is depicted in Fig. 10 (adapted from [200]). The experiment was conducted in an office with dimensions 2.6×5.6×2.7 m^3 (width, length, height), where the transmitter moved but the receiver remained stationary. The transmitter was initially placed at a distance of approximately 0.4 m from the receiver, and it was then moved 2 m to a distance of about 2.4 m away from the receiver before stopping. Afterwards, it was moved two more meters to about 4.4 m and stopped again. Finally, it was moved back following the same two steps. These movements can be observed in Fig. 10. The curves at around 10 and 20 s with negative velocity correspond to the two instances where the transmitter was moved away from the receiver. The next two instances (positive velocities) are caused by the transmitter being moved back in two steps. This velocity curve is accurate enough to be integrated to estimate the distance between transmitter and receiver, given an initial distance between them [200]. This property could potentially be used to enhance other forms of position information.

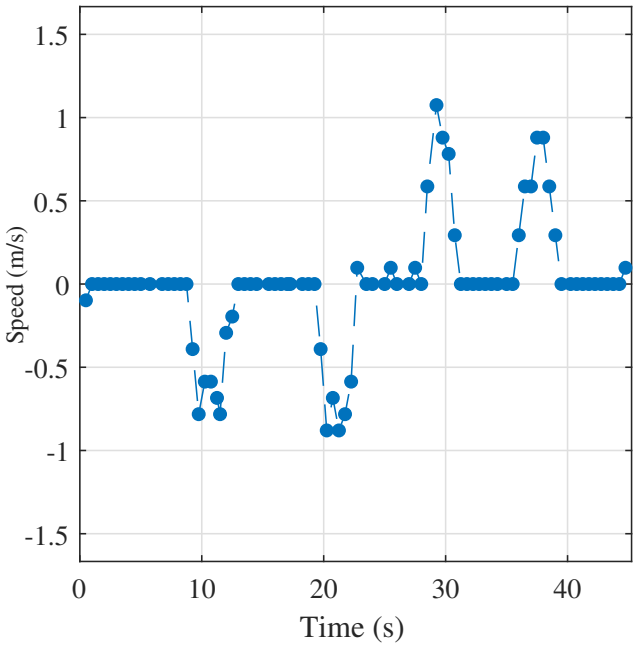


Figure 10: Measured velocity component along ultrasound beam, where negative values mean "away from transmitter" (adapted from [200]).

Next, different types of systems based on RSS will be reviewed.

A. *Binary signal strength*

These ultrasound location systems give a binary "yes/no" decision, corresponding to the presence or absence of signal, being the simplest system based on RSS. The utility of such a simple concept resides on the difference on how RF and ultrasound waves propagate indoors, as illustrated in Fig. 11. This figure shows the layout of an institution floor with many rooms, with two RF sources on the left side of the floor, in blue color, and two ultrasound sources (US) on the right side, in green. It can be seen how the RF signals spread out despite the walls, floor or ceiling. In this kind of environments, RF propagation models need to consider spreading, reflections, and attenuation in the walls, ceiling or floor [201]. However, the ultrasound signals are confined in the transmitter rooms, with the exception of some potential leakage through an open door. Therefore, the ultrasound source will be a reliable and simple indicator of the presence or absence of the transmitter in a room. It is worth noting that root mean square positioning

errors do not necessarily correlate well with the user experience. For example, an error of 1 m within the room usually does not have much consequence for the user, but on the other hand, an error of 1 m that locates the user in a different room, or floor, might have a considerable impact on the user experience. This is where the binary ultrasound positioning system finds its role and usage.

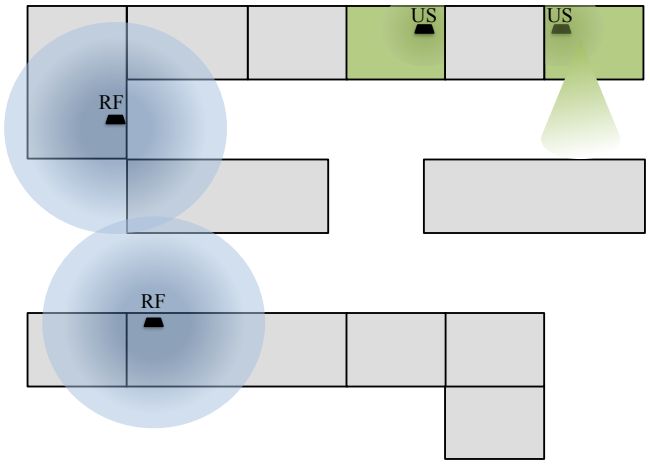


Figure 11: Indoor map showing an institution floor with many rooms. The ultrasound signals on the right (green color) are confined to the transmitter rooms, whereas the RF signals on the left (blue color) spread out despite the walls.

In practice, binary systems have two main applications: 1) They can be used as stand-alone systems, placing one node per room and simply outputting the room in which the tracked device is located. An example of this type of system, developed by Sonitor Technologies AS, is given in Section V. 2) These systems can be also used to resolve the aforementioned room ambiguities of RF-based positioning systems. From an economic point of view, WLAN-based positioning systems that leverage the existing infrastructure are very attractive. However, the performance of these systems seem to be limited in practice, providing median errors around 3 m, a 97th percentile errors around 10 m, and it seems that these limitations are fundamental [202], [203]. More complex environmental models might improve the performance, but large outliers can still appear resulting in errors that can locate the node in adjacent rooms or floors. One possible solution to solve this ambiguity is to add an infrared or an ultrasound positioning system. Whereas infrared has similar room-confinement properties as ultrasound, it is affected by fluorescent lamps and sunlight, and cannot cope so well with the lack of LOS conditions. Ultrasound systems, on the other hand, can do binary position detection based entirely on reflections, and therefore they seem to be preferable.

The system from the University of Geneva consisted of one loudspeaker per room, transmitting a pilot signal and a unique identifier [204]. Both parts were modulated using chirps at different frequencies. Upon identification of the pilot signal by a matched filter, the receiver could decode the identifier to obtain a binary signature, which allowed to identify the room. The system was robust to collisions by having different transmission periods at each loudspeaker. When multiple pilots are detected from different rooms, the system selected the one with highest energy. Experiments were conducted in 20 points spread over two rooms and a corridor. The system was tested with closed and open doors, using a Samsung Galaxy S5 as a receiver, with an app doing the processing. With closed doors, correct identification of the room was achieved for 100% of the trials. This performance slightly decreased to 99% with open doors.

B. Relative signal strength

After a binary signal strength decision, the next step up in complexity is to compare amplitudes. Using the knowledge that amplitude levels fall monotonically with range, these comparisons are relative and no propagation model is really required. These systems can provide better resolution than those based on binary signal strength.

In [199], the ultrasonic signal was received at several stationary base stations. The comparison of the received levels enabled the detection of the room section in which the tag was located. And in [205], the transmitter array depicted in Fig. 12 was used to transmit beams in different directions. Each beam was transmitted with a different FSK-modulated ID, and at different times. The receiver tag demodulated the beam ID and measured the RSS. The comparison of the received values for each beam allowed to figure out in which sector the receiver was placed. A possible application of such a system is bed-level resolution in a hospital.

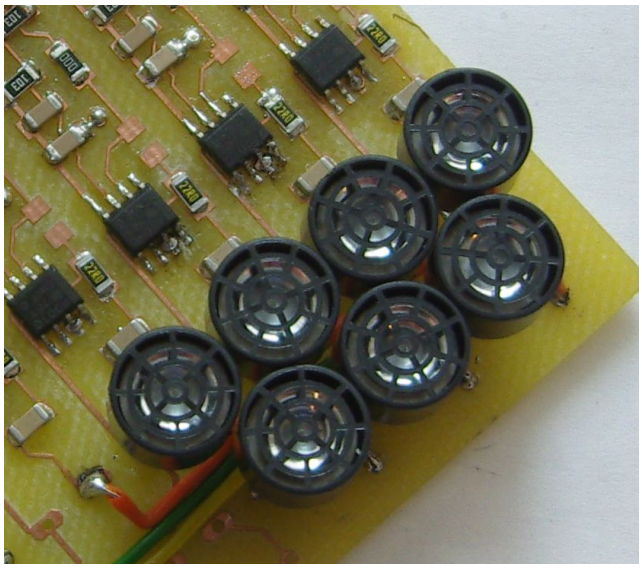


Figure 12: Ultrasound transmitter array with beamwidth of about 15° for sending unique coded beams into individual sectors [205].

C. Absolute signal strength

A propagation model is needed to use the actual signal strength value and to go beyond binary detection or relative signal strength comparisons. As long as LOS conditions apply,

spherical propagation combined with absorption is an accurate model and (4) describes the scenario for the propagation (or path) loss.

$$PL = 20\log(R/R_0) + \alpha R \quad (\text{dB}), \tag{4}$$

where R is the range between emitter and receiver, R_0 is a reference range, typically 1 m, and α is the absorption coefficient.

In [206], RSS values were used to calculate the distances between nodes in a network. The main novelty of this work by Shibaura Institute of Technology and the University of Tokyo, was the iterative multi-lateration approach to reduce the initial number of reference nodes with known positions. Five bidirectional transducers were used in each node to extend the coverage area in all directions, and the network was synchronized with an RF link, which was also used to exchange information to perform the location computation. A master node initiated the process, and commanded the transmitter nodes to send their ultrasonic pulses by means of a TDMA protocol. Both master and transmitter nodes knew their positions. When a receiver node measured enough distances, it performed multi-lateration and could become a transmitter node with known location in next steps, aiding in the localization process of other receiver nodes. The system included different error mitigation algorithms, such a priority-based node selection algorithm to minimize the accumulation of errors in the different steps, and an NLOS rejection algorithm based on the consistency of the received RSS values. The performance was evaluated in a small office environment, where 24 nodes were deployed. When using 4 of them as reference nodes, 95% of the errors were below 28.3 cm. This value decreased to 16.8 cm when using 6 nodes as references.

A star network topology was used by means of battery powered TELIAMADE nodes in [207]. One network coordinator configured and controlled the network, whereas a set of end-nodes were deployed in the ceiling, and one end-node was acting as a receiver. The network was connected by a Zigbee protocol. At the receiver, the ultrasonic signal was amplified, band-pass filtered and digitized, and the RSS values were obtained between pairs of nodes formed by the transmitters and the receiver, which were converted to distances in order to apply a lateration algorithm. In order to get good absolute calibration of amplitudes, special measures had to be taken into account, such as including the orientation of the transducers, and compensating for changes in battery voltage. In experiments using four transducers and a receiver placed at 10 different positions and heights, the system provided 95% of location errors below 19.5 cm.

Table VI gathers the characteristics and results from the previous systems, where accuracy refers to the maximum error considering 95% of the measurements, when a numeric value is given. A dash (–) means that the information was not available.

IV. Acoustic spectrum and other observables

Although TOF and RSS have been the most common observables exploited by positioning systems, there are other alternatives. In the last decade, the increase in computing power and the popularization of machine learning algorithms allowed to leverage a different observable, the acoustic spectrum, typically followed by a fingerprinting algorithm for localization [89]. This approach is specially interesting to deliver infrastructure-free positioning systems, where the user is assumed to carry a smart device with an embedded microphone. Therefore, the systems based on this approach have the potential to reduce considerably the deployment and operation costs, as no acoustic beacons would be required, only periodic updates of the fingerprints.

One of the first positioning systems based on the acoustic spectrum was the Batphone [208]. It was delivered as an application that provided localization support based on a combination of acoustic and Wi-Fi fingerprinting, being thus a purely passive infrastructure-free system (in the acoustic sense). The acoustic fingerprinting operation used the bandwidth from 0 to 7 kHz after the removal of transient sounds. In order to obtain the location, the smartphone's microphone recorded a signal of 10 s and calculated the spectrum next. It was then compared to the available fingerprints by means of the vector Euclidean distance and a nearest-neighbor classification method. An iPod Touch was used as the device in the experiments, and it was tested in a set of 43 different rooms. Using the acoustic spectrum by itself provided around 60% of correct detections, a performance that went up to around 70% when it was combined with the Wi-Fi fingerprinting.

Imperial College London presented a system to recognize rooms for forensic audio applications [209]. To classify the rooms they used the logarithms of the 1/3-octave band reverberation times from the room impulse responses. Experiments were conducted by analyzing 22 impulse responses from 22 rooms, applying a leave-one-out cross-validation procedure. The rooms were correctly identified in 96.1% of the trials.

EchoTag was presented as a system designed to recognize specific locations, or tags, by using a smartphone, in order to automatically activate certain functionalities on it, such as silent-mode, a timer, or a music player [210]. The system was developed as an application for iOS and Android devices, and it used the integrated speakers and microphones, thus being also infrastructure-free. It first used the information from the Wi-Fi Service Set Identifier (SSID) to check that the device was located in the correct room. It then checked the device's tilt, to make sure that the orientation fit some of the recorded tags. Only when the first two conditions matched at least one recorded tag, the acoustic sensing was activated. The device transmitted then a series of chirps distributed in four different frequency bands, which were used to sense the environment around 1 m from the device, approximately. The spectra of the received chirp echoes were evaluated to differentiate between the different locations by using a Support Vector Machine classifier, based on how different multipath conditions affected the spectrum. Using a 30 min dataset for the fingerprinting, EchoTag was able to distinguish 20 different tags in a home environment with a 95% accuracy, where these tags could also have a resolution of 1 cm and 30°. This accuracy fell after one week in the home set-up to 56%, due to changes in the nearby environment of the tags. However, in a lab environment the

Table VI: Summary of RSS systems.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
Univ. of Geneva	2016	20-21.5	No	≈ 0.2	Room-level	23.4-46.8	Yes	[204]
Univ. of Oslo	2010	40	No	1	Sub-room-level	22.4	No	[205]
Shibaura Inst. of Tech. Univ. of Tokyo	2004	40	Yes	0.5-1	16.8-28.3	12.96	No	[206]
Univ. of Granada Univ. of Oslo	2012	40	Yes	–	19.5	$\approx 0.5^{\dagger}$	No	[207]

† Area of the grid.

accuracy was still higher than 90% after a week.

The University of Queensland and CSIRO developed the Acoustic Landmark Locator [211] as a phone application. For each landmark to be considered (rooms and corridors), different measurements of the acoustic spectrum were taken with a mobile phone, and were then fed to an Encog neural network algorithm. These data were used to train the neural network, and the results were saved in the application. After being launched by a user, the phone application captured 1 s audio measurements, calculated the spectrum, and classified the room. Experimental tests were conducted at five different landmarks, with different levels of occupancy. Corridors provided the lowest performance, being correctly detected in 71% of measurements, whereas the correct detection of the room varied between 90 and 99%.

The system from the University of Peradeniya and the Sri Lanka Institute of Information Technology also employed a transmission/reception operation [212]. The transmitted signal consisted of a square wave of 6 kHz frequency and a duration of 50 ms, and the received signals were recorded. Different locations inside a room were evaluated. This system provided the best results when using autoregressive modeling on the power spectral density of the received signals. In that case, 60% of the estimated positions had an error below 1 m when using the average k-nearest neighbor method, considering two neighbors.

SoundSignature provided room-level localization based on selected features extracted from the acoustic spectrum [213]. During the offline stage, transient sounds were removed from the recorded data, several features were extracted and a Sequential Forward Feature Selection algorithm was used to reduce the total number of features. The selected features were all logarithms of the acoustic power at frequencies below 2 kHz. These features were then used to train a Support Vector Machine classifier. The effect of the smartphone position was tested on a small subset of four rooms, indicating that holding the smartphone on the hand, or having it in the pocket provided accuracies below 51%. Excluding these conditions, the difference in the results obtained with different smartphones was small. Two bigger datasets recorded on two different days in 16 rooms were used to analyze the performance. By using a 10-fold cross-validation on the first dataset, the accuracy was 90.28%, but this value dropped to 48.08% when using the dataset from one day to validate the data from the other day. This system could also use a secondary algorithm called SoundSimilarity, which compared the recordings of two different users by a cross-correlation, which could indicate that they were in the same room, providing a correction fix to the user with smaller confidence.

The prototype from Chiba University was based on the evaluation of environmental ultrasounds on the acoustic spectrogram by a fingerprinting approach [214]. Twenty measurement points were defined in a room, and for each, four different orientations were considered. Recordings with a duration of 2 s were used to calculate the spectrogram. Spectrograms were paired to position and orientation information, and data were collected to train a convolutional neural network that provided a non-linear regression model of the positions. Ignoring the orientation information, the obtained RMSE oscillated between 1 and 1.5 m, whereas the error was smaller when considering unidirectional datasets, indicating orientation of the microphone could not be properly obtained.

The sound absorption caused by the target to locate was exploited in the system developed by Mid Sweden University and the University of Salerno [215]. A subtraction method was employed, where the spectrum obtained when an object is present in the room underwent a smoothing operation, being subtracted next from the smoothed spectrum from an empty room, creating a fingerprint. In the experiments, nine different positions were considered inside the room, and a person was the target to localize. The recorded measurements at different positions were compared to the nine available fingerprints. As long as the person stayed within 29.9 cm from the measurement point where the fingerprint was taken, the system was able to obtain the correct position.

More recently, a combination of acoustic spectrum and RSS was used to calculate 2D positions using a single speaker [216]. The variation of the acoustic spectrum of a received chirp signal with the angle between the speaker and the receiver was used to estimate the DOA, based on a previously acquired database of measurements at different angles. RSS provided a range estimation based on a simple attenuation model. Experimental tests were conducted in a laboratory, considering 14 points at different ranges and angles. Combining all the estimations, 90% of the measurements had errors below 26.5 cm.

Finally, there are a few systems that are not based on TOF, RSS or the acoustic spectrum. For example, the system from

the University of Passau used the information obtained both from vibrations and from acoustic signals emitted by a device, in order to obtain a symbolic location of the target [217]. During the vibration phase, the device recorded the sound and the acceleration, and a set of features for each type of signals was extracted. During the sound sampling stage, 8 pulses at different audible frequencies were transmitted from the device speaker, and again a set of features were extracted. Classification of features is done with a C4.5 classifier algorithm, and a Bayesian Belief Integrator fused the three classifiers together. Three different rooms were evaluated in the experiments, considering a total of 35 different locations. Thirty measurements were performed at each of them, where 10 were used to train the classifiers, and 20 for the test set. According to the first result of the classification and assuming that the room is known, the accuracy of the location was between 90-94%. When the room was not known, the accuracy dropped to 78%.

Another example was Swadloon, which leveraged the Doppler shift experienced by a smartphone on the pulses sent by a set of fixed transmitters [218]. The phase and frequency shifts were measured by a Phase-Locked Loop (PLL), obtaining the relative displacement and velocity to the transmitters. This information was combined with the velocity from the internal inertial navigation system to calculate the direction to the transmitters by linear regression. Each transmitter used a different frequency band to avoid collisions. By shaking the smartphone to induce a Doppler shift, a static user could obtain the initial position, and be tracked afterwards when following a trajectory. Under static conditions, positioning errors were below 67 cm for 90% of the measurements at different points in an empty room, and below 1.23 m for 90% of the measurements in a big office space. For a 50 m trajectory, the errors were below 40 cm.

Table VII summarizes the characteristics and results from these systems based on the acoustic spectrum and other observables. Here, accuracy is given as the percentage of correct detections, with a note providing more details. A dash (–) indicates that the information was not available.

V. Commercial systems

This section reviews commercial positioning systems. Most of them do not provide enough information about their signal design, operation, and performance under real tests, so they cannot be readily compared to the previously reviewed systems, even if using the same observable to obtain the position. At the same time, some of them were, or still are available for the general public to purchase, so they are highly relevant to the indoor positioning field. Therefore, they have been gathered into their own section.

One early example came from Savi Technology, Inc. [219]. It consisted of a set of locators and tags attached to items, all synchronized with a RF link. The locators contained two ultrasonic transmitters, and the range to both transmitters could be calculated from the respective TOF. This setup also allowed to obtain bearing information from trigonometry, as the distance between the two transmitters in the locators is known.

The system from Netmor Ltd. consisted of a set of beacons deployed in the environment, which listened to the transmissions from moving units, attached to the targets [220]. Each unit could be differentiated by using a unique frequency, and the whole system was synchronized via RF, coordinated by a central unit, which commanded the start of the positioning cycle.

The Bat system [221] evolved from the ORL system [129], after the acquisition of Olivetti & Oracle Research Lab by AT&T Corporation. It was deployed in the three-floor office building of AT&T Laboratories in Cambridge, UK, and the authors reported tracking 200 users with 750 receivers units and three radio cells, with a claimed accuracy of 3 cm for 95% of the measurements.

Sonitor Technologies AS developed a positioning system for asset tracking based on binary signal strength [199], [222]. This system has been tested against more conventional RF-based systems with good results in a clinical setting [223], [224], and it consisted of wearable ultrasound tags and fixed ultrasound detectors, with a claimed communication range between 10 and 20 m. It was a stand-alone system with one node per room simply outputting which room the tracked device is located in, as introduced in Section III-A. The tags transmitted their unique ID, which were received by a detector, mounted on a room wall, or the ceiling. Multiple access was resolved by means of a Carrier-Sense Multiple Access (CSMA) protocol, where the transmitters checked the channel occupancy before sending their signal. All data from the detectors were collected in a central server, which could be accessed by a client terminal. This system later evolved into the current Sonitor Sense [225], which works in the opposite way. In this system, there is one transmitter with unique ID per room, where it can create up to four different areas, obtaining sub-room level accuracy. Each user carries a receiver tag, which also has a unique ID. Upon reception of the transmitter signals, the tags send back their location information and ID to a server via Wi-Fi.

SmartLocus was a prototype developed by Hewlett-Packard Laboratories [226]. In this system, tags were also attached to the nodes to localize. Some of these nodes were defined as infrastructure nodes, and they were expected to be mostly located at fixed positions. Other nodes were mobile, and therefore they could move around the environment, being tracked by the system. If they stop moving, they could become an infrastructure node and help other nodes to localize themselves. A mobile node would look for its neighbors and obtain ranging information from them. This was done upon request via RF and, after that, infrastructure nodes replied back with their own location information via RF and an ultrasonic pulse, following a random back-off strategy to minimize collisions. Once the mobile node obtained its position, it could send it to the back-end application.

NEC (China) Corp. presented a prototype called AUITS [227]. This system was based on a unit of receivers with structural topology, which provided easier deployment and calibration, and a tag to be attached to the target to be located. Synchronization was obtained by RF, and outlier rejection was performed on the TOF according to triangle inequalities

Table VII: Summary of systems based on the acoustic spectrum and other observables.

Name / Developers	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (%)	Area (m^2)	Privacy	Reference
Batphone	2011	0-7	No	–	60-70[1]	–	Yes	[208]
Imperial College London	2013	0.16-3.15	No	–	96.1[1]	29-9500[†]	Yes	[209]
EchoTag	2015	11-22	No	–	40-98[2]	–	Yes	[210]
Univ. of Queensland CSIRO Univ. of Wollongong	2015	0-11	No	–	71-99[1]	75-225	Yes	[211]
Univ. of Peradeniya Sri Lanka Inst. of Tech.	2017	6	–	–	60[3]	42	Yes	[212]
SoundSignature	2018	0-2	No	≈ 0.2	77.9-90.3[1]	–	Yes[‡]	[213]
Chiba Univ.	2019	20-125	No	–	1-1.5[4]	28	Yes	[214]
Mid Sweden Univ. Univ. of Salerno	2020	5-20	No	–	Sub-room level[5]	11.2	No	[215]
Hokkaido Univ. The Univ. of Tokyo Nat. Inst. of Informatics	2021	6-24	No	–	26.5[6]	1.5	Yes	[216]
Univ. of Passau	2007	0.5-4	No	–	78-94[7]	–	Yes	[217]
Swadloon	2015	17-19.5	No	4-20	67-123[6]	180	Yes	[218]

[1] Room-level localization.
[2] Sub-room level localization: can differentiate between tags separated 1 cm.
[3] Sub-room level localization: correct detection inside a radius of 1 m.
[4] RMSE in m.
[5] Accuracy of 100% if target was closer than 30 cm from reference points.
[6] Maximum errors considering 90% of the measurements.
[7] Sub-room level localization: correct detection at different abstract locations.
[†] Room volume.
[‡] When not using SoundSimilarity.

between the tag and two receivers. Upon reception of the ultrasonic signal, the tag position was calculated by trilateration using the three receivers with maximum separation between them. This position could be forwarded to a central server for location services. Experiments were performed by employing a unit with 7 receivers. Positions under static conditions were calculated at 15 points inside a room, whereas a circular trajectory was followed to test the system under dynamic conditions. The obtained errors were below 8 and 22 cm, depending on the distance to the unit center, considering 90% of the measurements.

The system from Celltek Electronics Pty Ltd. consisted of a set of RF transceivers and ultrasonic transmitters deployed in a certain space, such as a shopping mall [228]. A tag provided to the user could synchronize via RF to the system, and it listened to the ultrasonic transmissions, which were frequency-modulated signals transmitted on a TDMA protocol. The tag could either calculate the position, or just the distances, and it sent either information to a computer for calculation, tracking, and displaying purposes. One interesting feature of this system was that it included a backup positioning method based on the RSS values of the RF signals, in case the ultrasonic signals could not be received. Inertial measurement units included in the tags could also be employed to travel between areas with no coverage, until receiving a position fix.

AeroScout Ltd. developed a WLAN system with wearable tags that included ultrasound receivers for room-level resolution [229]. In this system for assets management, ultrasound transmitter nodes must be placed in each room where location ambiguity is particularly important to resolve. The transmitted unique ID was received by the wearable tag, which reported back the detected room-ID to a server via the WLAN. The server could estimate then the location of the tag. AeroScout was later acquired by Stanley Healthcare Solutions [230], and the system is still available. They also announced a partnership with Sonitor Technologies AS to use Sonitor Sense [225] in the AeroScout RTLS for the healthcare industry [231].

Intelligent Sciences Ltd. designed a system that used one base station located in the ceiling, and a transponder as receiver [232]. Upon reception of an ultrasonic pulse from the base station, the transponder sent back another pulse, which allowed to calculate the Round-trip Time-Of-Flight (RTOF) to several receivers in the base station. In that way, both range and angle to the transponder from the base station could be calculated. Alternatively, a synchronized version of the system allowed the transponder to act as an emitter and send its pulse when instructed by the synchronization signal, calculating the TOF at the base station.

The system from KT Corporation consisted of a set of synchronized beacons (satellites), which transmitted sequentially an ultrasonic signal [233]. These signals were received by a device, which was also synchronized to the beacons, so it

could calculate the TOF. The device could also get the beacons position by accessing a server that contained such information.

An earlier system from CenTrak, Inc. consisted of an RF base station to provide synchronization, and several ultrasonic stations that transmitted the station ID following a TDMA protocol determined by the RF station [234]. A portable device carried by a user was synchronized to the system by receiving the RF information, and upon detection of the base station code, it transmitted back both the station ID and its own device ID to a server via RF. The position was then associated to the detected ultrasonic station. A recent version uses infrared base stations to synchronize fixed ultrasonic transmitters with a portable tag [235]. After calculating the distances to the transmitters, the tag can either calculate its position by trilateration, or just calculate the distances to the transmitters. In any case, position or distances are sent to the server for displaying purposes.

The positioning system from Marvelmind Robotics [236] consists of a set of beacons deployed on the walls or ceiling, and a mobile beacon attached to the target. All beacons are synchronized with RF, and the system can work both in a privacy-oriented or centralized version. Additionally, the positions from ultrasound can be fused with an IMU. When the mobile beacon is listening (privacy-oriented), each transmitting beacon uses a different frequency, from 19 to 45 kHz, so they can be distinguished. This system was used in [237] to locate a drone that conducted a structural health monitoring test in a room, using four static beacons and a mobile beacon on the drone. The position obtained with the ultrasonic system was fused with the drone's inertial navigation sensor. Different tests were performed with two drones, while hovering or following a rectangular trajectory of 11 m^2. In the hovering tests, absolute errors were between 10.2 and 15.8 cm for the 2D position of the drone. In the trajectory test, the maximum absolute error calculated at each corner of the trajectory was approximately 18 cm.

The system from Hexamite considered a set of transmitters and receivers, synchronized by an RF link [238]. After calculating the TOF and ranges to the transmitters, the receivers sent this information via RF to a central station, equipped with an RF antenna connected via Universal Serial Bus (USB). This central station could then calculate the position and orientation of the target.

In the system from Certis Cisco Security Pte Ltd. [239], the mobile device captures sound recordings, which are transferred to a central unit. This unit processes the audio to evaluate the peaks and valleys of the spectrum and compares those to a database in a fingerprinting operation. The recordings can contain information from an acoustic transmission, as well as the background noise, and can potentially be aided by camera pictures, and also by tags communicating via bluetooth or near-field protocols.

One of the latest systems has been developed by Forkbeard Technologies AS [240]. It provides a positioning system for smart devices consisting of a set of beacons deployed in the environment, and it is available for Android and iOS users as an app (Forkbeard Lyra). It combines a communication device with ultrasonic signals to calculate the TOF at the smart device upon entering the area covered by the beacons. Each beacon can transmit a unique QPSK modulated Complementary Code Keying (CCK) in a specific time slot to avoid collisions and reduce reverberation. These signals are received by the smart device, which can identify the beacon based on the code and transmission slot. Its decoder can then estimate both range and radial velocity to each beacon, and can compensate for multipath. By combining both estimates into a Kalman-like algorithm [241], it obtains the "Doppler stabilized" ranges. These ranges are then used in a multi-lateration process to estimate the position [242]. This system was recently the subject of a characterization study [243]. When using 10 beacons, and measuring 100 positions in each of a 73-point grid by placing a smart phone on the floor, 80% of the measurements had errors below 44 cm. Dynamic conditions were also tested and compared to a motion capture system. When repeating five times a 42-meter-long walking trajectory while carrying the phone in the hand, this error level increased to 88 cm. For the worst case scenario, repeating the same trajectory five times while carrying the phone in the pocket and with obstacles inside the room to recreate NLOS conditions, 80% of the obtained positions had errors below 1.02 m. The Forkbeard system was later used together with a continuous output dead-reckoning system that considered a sine-wave approximation on the step detection algorithm [244]. This step detection algorithm was combined with heading estimation and the ultrasonic measurements into an EKF, and the system was tested by following a trajectory in a room with NLOS conditions, obtaining a smoother and continuous trajectory, compared to the ultrasonic system alone.

Table VIII summarizes the main parameters from these commercial systems. The accuracy and area values are those reported by the companies, unless specified in the table. A dash (–) means that the information was not available, and a hashtag (♯) next to the company name means that the product is not available for sale, to the best of our knowledge.

VI. Link Budgets

Most papers on acoustic positioning deal with accuracy in the spatial domain measured in centimeters and millimeters. But the experience of many was reflected in what [245] said when they stated that they had yet to see a purely acoustic tracker that doesn't "go berserk" when you jingle your keys. Very few deal with how robust the systems are with respect to noise, or try to predict their useful range under various conditions. Therefore, in this chapter we consider the link budgets to analyze the expected useful range of different acoustic positioning systems.

This useful range is given by the communication link between emitter and receiver, which is similar to that of other one-way systems, such as satellite links or passive sonars. Therefore, the range can be obtained from the passive sonar equation [193]:

$$SL + PG - PL - NL > DT, \tag{5}$$

where here SL is the source level in dB SPL (sound pressure level relative to 20 μPa) at a reference range, usually $R_0 =$

Table VIII: Summary of commercial systems.

Company	Year	Frequency (kHz)	Synch.	Update rate (Hz)	Accuracy (cm)	Area (m^2)	Privacy	Reference
Savi Tech. Inc.♯	1996	–	Yes	–	–	–	No	[219]
Netmor Ltd.♯	2000	–	Yes	–	–	–	No	[220]
AT&T Lab.♯	2001	–	Yes	50	3[1]	≈ 930†	No	[221]
Sonitor Tech. AS♯	2002	25	No	–	Room-level	10-20‡	No	[199], [222]
Sonitor Tech. AS	2014	40	Yes	1	Sub-room level	–	No	[225]
Hewlett-Packard Lab.♯	2005	40	Yes	1	–	–	No	[226]
NEC (China) Co. Ltd.♯	2008	40	Yes	–	≈ 8-22	15	No	[227]
Celltek Electronics Pty Ltd.♯	2009	–	Yes	–	–	–	No	[228]
AeroScout Ltd.	2010	–	No	–	Room-level	–	No	[229]
Intelligent Sciences Ltd.♯	2010	40	Optional	–	–	–	No	[232]
KT Corp.♯	2010	–	Yes	–	–	–	Yes	[233]
CenTrak, Inc.	2013	–	Yes	–	–	–	No	[234]
Marvelmind Robotics	2014	19-45	Yes	25	10.2-18[2,*]	170*	Optional	[236], [237]
Hexamite♯	2016	40	Yes	20	0.9[3]	–	No	[238]
Certis Cisco Security Pte Ltd.♯	2020	0.02-7	No	–	Room-level	–	No	[239]
Forkbeard Tech.	2020	20-21	Yes	16	44-102[3,*]	150*	Yes	[240], [243]

[1] Maximum error for 95% of the measurements.
[2] Absolute error at selected waypoints.
[3] Absolute error.
[4] Maximum error for 80% of the measurements.
♯ Discontinued.
† Total area over three floors of the building.
‡ Maximum range.
* Results from independent experiments.

1 m; PG is a processing gain, which is 0 dB for the simplest systems; PL is the propagation loss; NL is the noise level; and DT is the detection threshold. Next, we briefly discuss these variables and give some typical values that will be used in the subsequent analysis.

The source level for a multi-element array and in the direction where the beam is steered will be $SL = SL_1 + 10 \log N_{tx}$, where SL_1 is the source level of each individual element, and N_{tx} is the number of elements [246]. The propagation loss consists of two terms, as given in (4). The absorption α (dB/m) is caused by relaxation processes in oxygen and nitrogen molecules, and is a function of temperature, humidity, and pressure [247]. Fig. 13 shows the absorption for 20°C and 1 atm. While commonly neglected at audio frequencies, it can be relevant at long ranges and high frequencies. At 40 kHz it takes values from $\alpha = 0.26$ dB/m for 0% relative humidity (RH), to a maximum of 1.33 dB/m for 55% RH, and then falls to a value of 1.1 dB/m at 100% RH.

At ultrasound frequencies, there are not many noise measurements which have been published, but some data can be found in [248] for an industrial environment: considering a 3 kHz measurement bandwidth, they give a level of 70-80 dB SPL in the range 20-60 kHz, where air tools may produce levels up to 100 dB SPL. We have considered here a relatively low level of 60 dB in our calculations, but it is worth noting that, in extreme cases, the noise level may actually be up to 40 dB higher. As for a low noise value, a background level at audio frequencies for a quiet library is about $NL = 40$ dB SPL, which is assumed to apply at ultrasound frequencies as well.

The critical design parameter that determines the influence of the noise is the processing bandwidth BW, as the noise level which should be used in (5) is:

$$NL = NL_0 + 10 \log(BW) \tag{6}$$

Then, the equivalent noise spectral density for $BW = 3$ kHz and $NL = 60$ dB SPL is $NL_0 = 60 - 10 \log 3000 \approx 25$ dB/Hz. Therefore, considering the variation in NL between the library and industry environment with air tools, there may be a variation in the noise spectral density of $NL_0 \in [5, 65]$. As we do not consider it to be realistic for ultrasound systems to work in the extreme cases such as in the vicinity of air

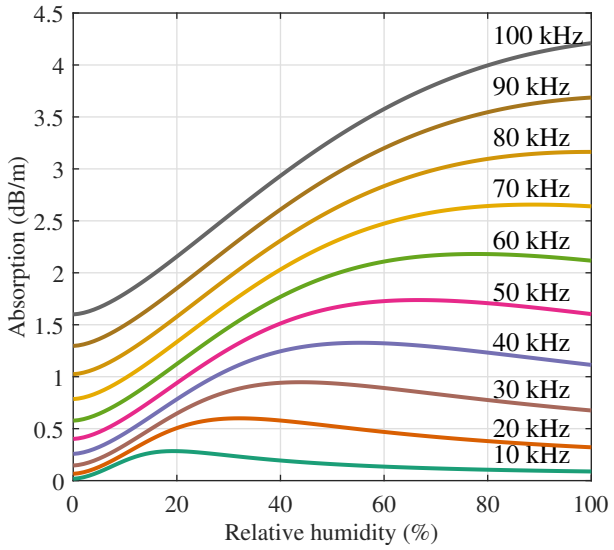


Figure 13: Absorption (in dB/m) as a function of frequency and humidity, at 1 atm pressure and 20°C temperature.

tools, we will do the analysis for $NL_0 = 25 \pm 20$ dB/Hz. This large variation is one of the factors that distinguishes ultrasonic systems from radio communications. It also explains why a system might easily fail in real life while working perfectly fine in a quiet lab over a long range.

The system analysis that follows has been done with the assumption of a center frequency of 40 kHz and a worst case absorption of $\alpha = 1.33$ dB/m. In optimal detection theory, a detection threshold from 15 to 20 dB depending on probabilities of detection and of false alarm is common [249]. Here the value is assumed to be $DT = 20$ dB as in [200].

A fairly standard transducer such as the Murata MA40S4S is capable of providing 120 dB SPL at $R_0 = 0.3$ m, which is equivalent to about 110 dB SPL at $R_0 = 1$ m [250]. This is when driven with a sinusoidal wave of 20 V_{p-p}. A 10 dB reduction in source level thus corresponds to driving it with about 3.2 V_{rms}. Other values of source level may be analyzed easily since the source and noise levels are simply subtracted from each other in the link budget. The ± 20 dB variation in noise level may therefore equally well be interpreted as a variation in source level. In this work, the source level is set to $SL = 100$ dB. This value also aligns with health and safety reports for ultrasound components above 20 kHz [251]. It is worth mentioning that maximum values are commonly given for occupational exposures, but on recent years maximum permissible levels are under scrutiny for public exposure [252].

Lastly, the processing gain PG will have two components, one due to pulse coding, PG_{code}, and one due to receiver array gain, which depends on the number of elements N_{rx}: $PG = PG_{code} + 5 \log N_{rx}$. The receiver array sums coherently for the signal and incoherently for the background noise, so that is why the dB calculation only has a 5 as multiplication factor.

A. Narrowband systems

For this case we have considered $PG = 0$ dB, since there is no coding, and $BW = 25$ Hz, which is far below the threshold of 10% the center frequency. The link budget for estimating range in this case was first given in [193], [200], and it is shown here in Fig. 14. The red-dashed line shows how the source level is attenuated with distance, considering only spherical spreading. The blue line also takes into account the absorption term into the attenuation, assuming a relative humidity of 55%. The lowest horizontal line in green color is the noise level, and above that is the line for the detection threshold, in purple. The maximum range is obtained when the power curve (blue line) crosses the detection threshold giving a range of about 14.4 m, marked with a black circle. The ± 20 dB variation of the background noise level is shown as an orange band around the detection threshold, and it represents the operating range of the system. It gives a range of slightly less than 6 m for the maximum noise level (top edge of the orange band crossing the power curve), and a range of about 26 m for the lowest noise level (bottom edge of the band crossing the power curve). This is about the same range as for the coded system in Fig. 16 as first shown in [253]. This is because the processing gain of the coding makes up for the reduction in range due to the increased bandwidth of the coded system.

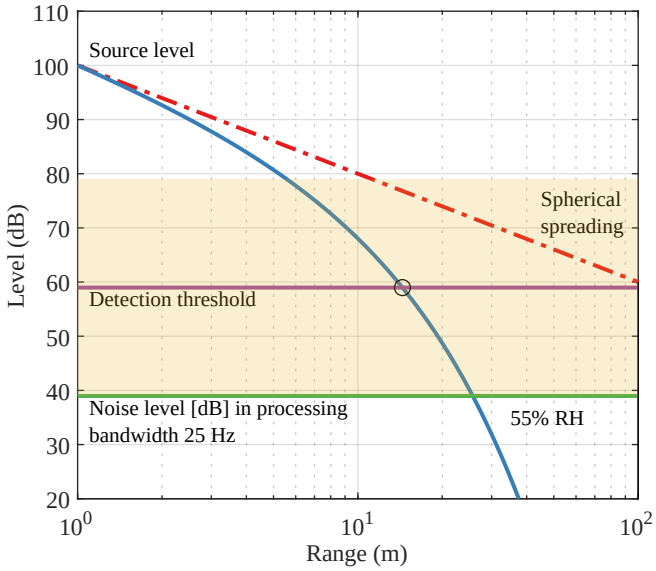


Figure 14: Link budget for a narrowband system with bandwidth 25 Hz, and $PG = 0$ dB.

B. Wideband uncoded systems

In a wideband system, a representative bandwidth for a 40 kHz system is 4 kHz, and since it is uncoded, $PG = 0$ dB. The link budget is depicted in Fig. 15. In this case the power curve crosses the detection threshold at a range of about 5 meters. Again, the orange band indicates the ± 20 dB variation in the background noise level. As the top edge of the band is over the power curve, one finds that at the maximum noise level the range is zero, i.e. the system will not work at all. On the other

hand it will give a range of about 15 meters for the lowest noise level.

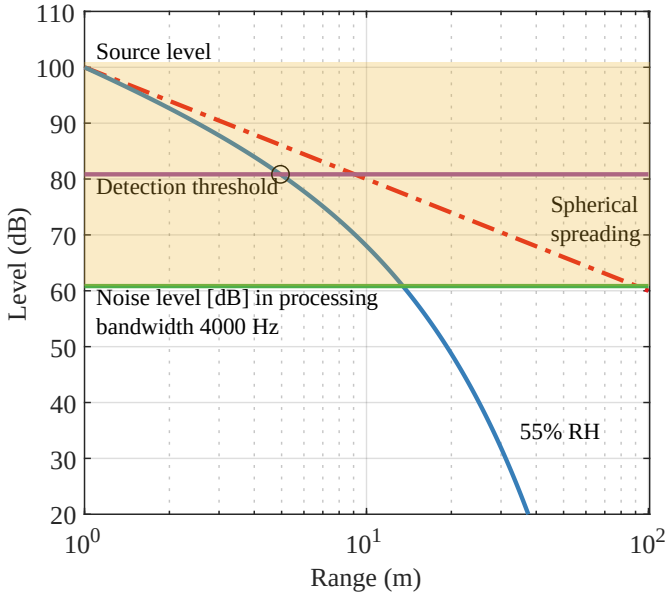


Figure 15: Link budget for a wideband uncoded system with bandwidth 4 kHz, and PG = 0 dB.

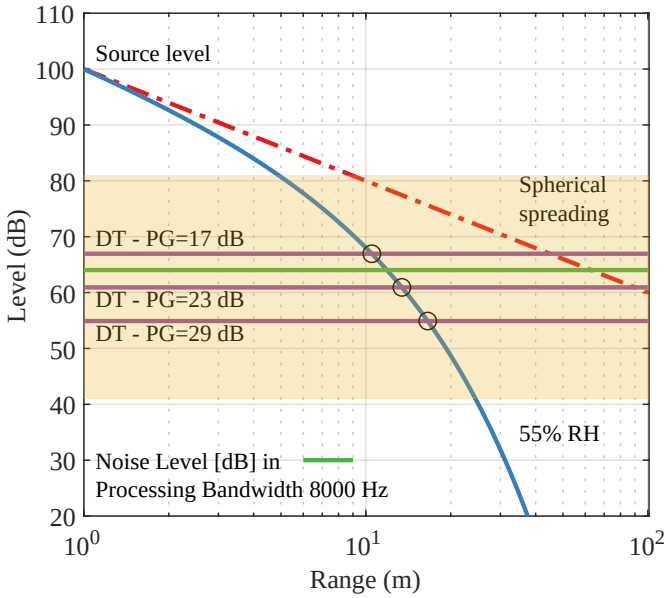


Figure 16: Link budget for a wideband coded system with bandwidth 8 kHz and processing gain PG of 17, 23 and 29 dB, corresponding to sequence lengths L of 256, 1024 and 4096 bits.

C. *Wideband coded systems*

Regarding the link budget calculation, and in comparison with the uncoded system previously studied in Fig. 15, the bandwidth BW is set to 8 kHz in this case, since a larger bandwidth is necessary with coded and modulated transmissions. The main difference is that now there is a processing gain PG due to the coding technique. The processing gain depends not only on the length L of the sequences used to encode ultrasounds, but also on the features of the modulation. However, since in most cases the effect of the modulation is negligible compared to the effect of the codes applied, the processing gain PG is considered here as a function only of the code length L, according to $PG = 10 \log L - 7$ dB. The loss of 7 dB is due to power control errors and threshold detection effects (see [253]).

The result is shown in Fig. 16 for the cases of sequences with lengths 256, 1024 and 4096 bits, which provide processing gains of 17, 23 and 29 dB, respectively. The maximum range is found when the power curve (blue) crosses the detection thresholds, giving a range of about 10, 13 and 17 m, for the three sequences. Considering the link budget for the 1024-bit sequence ($PG = 23$ dB), with the variation of the background noise level of ± 20 dB represented by the orange band, the system provides a range of 5 m for the maximum noise level, and a range of about 26 m for the lowest noise level.

VII. Lessons Learned and Future Challenges

Traditionally, pioneering acoustic positioning systems have been characterized by some persistent drawbacks that have hindered their development and deployment in numerous fields. The most relevant ones may be a low accuracy, a reduced update rate and a high sensitivity to in-band noise. These inconveniences are already present in the most basic configurations, particularly in those based on narrowband pulse-echo structure. Nevertheless, they are often considered in some applications where simplicity is a must for different reasons.

In order to overcome these issues, the wideband approach, uncoded and coded, has been proposed and evolved during the last decades. In this case, the resulting solutions provide a higher accuracy and robustness to noise. Particularly, the use of different sequences with suitable correlation properties has allowed the accuracy to be boosted to the range of few centimeters, while improving the coverage achieved by the beacons, due to the gain provided by the matched filtering used in the detection. On the other hand, these spread-spectrum techniques often involve longer transmission times, which consequently constrain the final update rates.

Furthermore, commercial transducers usually offer reduced bandwidths, in the range of a few kHz, which implies an important constraint to be tackled when designing the acoustic signal to be transmitted to the environment. This consideration also applies to smart devices, which favored an increased research interest as smartphones became widespread, since those devices could be used instead of dedicated hardware. Positioning systems based on smart devices have experienced some issues related to computing power and low resolution due to the limited bandwidth and sampling frequency available in standard COTS smart devices, which were not originally intended for indoor positioning. New hardware that expands the acoustic frequencies to ultrasound could provide better results, although this could also limit the coverage area. This is a significant aspect that still remains to be solved.

Another limitation typically associated to acoustic positioning systems is the Doppler shift that appears in the propagation of the ultrasonic wave, due to the movement of any element in the system, either transmitter or receiver. Due to the relative low propagation speed of sound in air, common vehicles and objects that move in our daily life might have an impact on the acoustic wave, which constrains the possible application of these systems to static scenarios or scenarios with a reduced capacity of movement. This has led to proposals of numerous algorithms and methods to firstly estimate this Doppler shift, and then compensate it. Nevertheless, these algorithms still have room for improvement, and they always involve a remarkable computational load at the signal detection. It is worth noting that this computational complexity is sometimes relegated as a secondary aspect, but it emerges at the final stages of experimental prototypes as an ultimate feature in order to achieve feasible solutions, ready to be developed and deployed by industries in many fields. In this way, it is necessary to come up with practicable methods that can be implemented in real time.

Likewise, the multipath effect is one of the most constraining factors for acoustic positioning systems. As has been previously mentioned, multipath becomes particularly problematic in NLOS cases, when a reflected path may be considered as the LOS echo and significantly degrades the position estimates and the final performance of the system. For that purpose, some works have proposed methods mainly focused on the mitigation of the multipath effect and the NLOS rejection, thus improving the accuracy. They consist of the successive interference cancellation, and, sometimes, they are also able to deal with the inter-symbol interference caused by the limited bandwidth of the receivers, as well as with the multiple-access interference originated from stronger emissions from nearby beacons (near-far effect). In some cases, these algorithms try to identify outliers in the position estimates in order to discard them as soon as possible. In this context, the installation of additional transmitters and/or receivers might provide the extra information necessary for a suitable distinction between LOS and NLOS echoes. Nevertheless, all these proposals have obtained a reduced success, mainly only in well-known and controlled environments, and they are still far from a suitable generalization that can cope up with multipath issues in any context. Furthermore, these methods are often iterative, which implies a high computational load. This aspect has already been mentioned before and constrains the achievement of suitable experimental prototypes. These two items have become the main challenge for handling multipath in acoustic positioning systems.

Another drawback is the distance ranging or coverage achieved in real environments. The covered area in acoustic systems is typically smaller than in those based on radiofrequency, due to the high attenuation of sound in air, and also the typically high directivity of ultrasound transducers. Although coded wideband systems have allowed the distance ranging to be increased, due to the process gain of the sequences involved, real experimental cases often require the deployment of numerous beacons to properly cover the volume under study. This results in a twofold issue: on one hand, acoustic positioning systems need the installation of specific infrastructure in the environment, which may increase costs; and, on the other hand, solutions often involve other sensory systems to complement the acoustic one. Different works have already been described that integrate other sensors, such as inertial units, RF modules, or odometrics. This integration usually involves a fusion method, where the Kalman filter and its different variations have been predominant. Anyway, both aspects, infrastructure installation and hybridization with other sensors, should be tackled in coming years to address and achieve a larger coverage with enough reliability and availability.

With regard to fingerprinting solutions, the need to characterize the room firstly by identifying its unique signature, implies the same restrictions that can be found in other technologies, such as WiFi positioning. In this context, acoustics is particularly sensitive to any change in the furniture distribution, and patterns should be obtained under controlled conditions (i.e., windows and doors closed) on a fine grid throughout the room. Afterwards, real cases where people can move in and out substantially modify the correlation between current measurements and templates, and degrade the final performance. This implies two immediate drawbacks: on one hand, the process of acquiring the templates is laborious and tedious, and they often require periodic updating for an optimal operation; on the other hand, NLOS situations and multipath become critical again, since significant errors may appear. At the same time, the advent of machine learning and more powerful hardware embedded in smart devices have favored the appearance of infrastructure-free systems. These systems usually do not need any dedicated infrastructure, since no beacons are required, and the measurement and processing stages can be run on an app installed in a smart device. Despite their current limitations, they have the potential to reduce the deployment and maintenance costs significantly. Combining this approach with a crowdsourcing philosophy could potentially help in this updating process.

Depending on the accuracy, there are two main areas of application in which acoustic positioning systems have reached enough maturity: 1) The positioning of objects in a single room within a limited range, such as navigation applications, either for people or robots. 2) Logistics or personnel tracking in large institutions such as hospitals. Room-level positioning can give enough accuracy in this case, where a binary received signal strength system can be suited. Good coverage and robustness are required in these applications, even when the range approach 10 m. If accuracies within 1-3 m are needed, then combining a binary received signal strength ultrasound system with a Wi-Fi localization system is well suited. Systems based on the entire echo structure, the acoustic spectrum or absolute signal strength are not mature enough yet to have reached commercial applications.

In general, we have observed that systems with similar design and coverage area typically provide similar accuracies. Variations are probably caused by the different hardware and experimental procedures, such as the number of grid points (which might or might not include non-line of sight measurements), use of accurate ground truth to measure trajectory

errors, or the figure of merits used to report such errors (e.g. mean square errors, absolute errors or maximum errors for a certain percentage of measurements). In this regard, some effort has been made in the last decade at the Indoor Positioning and Indoor Navigation conference (IPIN) [254], and the Microsoft Indoor Localization Competition (IPSN) [255]. Both events provide a common setup where different positioning systems can be tested. A common benchmark would indeed be a promising way to compare different systems.

Based on the trends observed in recent years in the development of location-based applications, we may identify several future challenges for local acoustic positioning systems. Firstly, the positioning of people in large environments, with strongly changing conditions in terms of furniture distribution and/or public flow, and which should also be easily scalable (extending its coverage should not require a complete redesign of the system). Secondly, the continuous positioning of vehicles with complex dynamics such as mini and micro unmanned aerial vehicles that are currently used in inspection tasks and whose rotors are a powerful source of high frequency acoustic noise. Thirdly, using speakers and microphones that support higher frequencies than those of current COTS smart devices would allow to leverage ultrasonic signals with wider bandwidth, which can further increase the system accuracy without the need of external dongles. Lastly, there is also a trend towards the development of location-based services that allow seamlessly transit between mixed indoor-outdoor environments, which cannot be ignored by acoustic technology, specially in areas with poor GPS coverage. The use of this type of acoustic positioning signals in an uncontrolled atmosphere, with the presence of meteorological phenomena such as temperature gradients, aerosols or turbulence, which all have a significant effect on the propagation of this type of mechanical signals, is undoubtedly one of the most fascinating challenges posed by this field of research.

Appendix A
Speed of sound variation

The speed of sound is not a constant but will vary with temperature and humidity, as well as pressure and frequency. The speed of sound varies approximately linearly with temperature in dry air:

$$c = \sqrt{\frac{\gamma k T}{m}} = 331.45\sqrt{1+\frac{t}{273.15}} \approx 331.45 + 0.607t \quad \text{(m/s)} \tag{7}$$

where γ is the adiabatic index, k is Bolzmann's constant, m is the average mass of a single molecule, T is the temperature in K, and t is the temperature in °C. The nominal value at 20°C is $c_{20} = 343.4$ m/s. The sensitivity to temperature variation can be found by computing $s = dx/dt/x \approx 0.607/343.4 \approx 1.8$ mm/m/°C. Thus, at a range of 5 m and with a temperature error of 2°C, the error will be 1.8 cm, and in a high precision system this needs to be compensated for. In [256] one of the first works that tried to compensate temperature drifts in a 2 m workspace for the positioning of a robot's end effector can be found. The authors modeled the range estimation noise caused by temperature drift and turbulence, and used it in a Kalman filter to compensate the ultrasonic measures for these effects.

Humidity causes changes in γ and in m which lead to an increase in sound speed with humidity. A change from 0 to 100% relative humidity at 20°C causes a sound speed change of less than 0.4% [257]. At a range of 5 meters this may lead to an error of 20 mm. If the range of variation of the humidity is small, changes due to humidity can often be neglected.

Eq. (7) does not show any variation with pressure. This is an approximation and in reality the speed of sound has a positive sensitivity of 1 ppm (m/s per kPa) around 20°C [258]. This can be neglected under normal conditions.

In an absorbing medium where the absorption varies differently from the square of the frequency, there will be dispersion, i.e. sound speed variation with frequency. The range of variation is less than 0.15 m/s, so it is usually assumed negligible [259], [260].

References

[1] U.S. Environmental Protection Agency, "Report to congress on indoor air quality: Volume 2. EPA/400/1-89/001C," 1989.

[2] World Health Organization, "Combined or multiple exposure to health stressors in indoor built environments," Available online at https://www.euro.who.int/__data/assets/pdf_file/0020/248600/Combined-or-multiple-exposure-to-health-stressors-in-indoor-built-environments.pdf, 2013, last accessed on: July 22, 2022.

[3] MarketsandMarkets, "Indoor Location Market worth $17.0 billion by 2025," Available online at https://www.marketsandmarkets.com/PressReleases/indoor-location.asp, 2020, last accessed on July 22, 2022.

[4] M. N. K. Boulos and G. Berry, "Real-time locating systems (RTLS) in healthcare: a condensed primer," *International Journal of Health Geographics*, vol. 11, no. Article number 25, pp. 1–8, 2012.

[5] F. Santoso and S. J. Redmond, "Indoor location-aware medical systems for smart homecare and telehealth monitoring: State-of-the-art," *Physiological Measurement*, vol. 36, no. 10, pp. 53–87, 2015.

[6] I. Hwang and Y. J. Jang, "Process mining to discover shoppers' pathways at a fashion retail store using a WiFi-base indoor positioning system," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 4, pp. 1786–1792, 2017.

[7] N. Ferracuti, C. Norscini, E. Frontoni, P. Gabellini, M. Paolanti, and V. Placidi, "A business application of RTLS technology in Intelligent Retail Environment: Defining the shopper's preferred path and its segmentation," *Journal of Retailing and Consumer Services*, vol. 47, pp. 184–194, 2019.

[8] S. Kohlbrecher, O. von Stryk, J. Meyer, and U. Klingauf, "A flexible and scalable SLAM system with full 3D motion estimation," in *Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robotics*, Kyoto, Japan, Nov. 2011, pp. 155–160.

[9] G. Glanzer, "Personal and first-responder positioning: State of the art and future trends," in *Proceedings of the Ubiquitous Positioning, Indoor Navigation, and Location Based Service conference*, Helsinki, Finland, 2012, pp. 1–7.

[10] A. Lim and K. Zhang, "A robust RFID-based method for precise indoor positioning," in *Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, Annecy, France, 2006, pp. 1189–1199.

[11] N. Jathe, M. Lütjen, and M. Freitag, "Indoor positioning in car parks by using Wi-Fi round-trip-time to support finished vehicle logistics on port terminals," in *Proceedings of the Conference on Manufacturing Modelling, Management and Control*, Berlin, Germany, 2019, pp. 857–862.

[12] I. Rubino, J. Xhembulla, A. Martina, A. Bottino, and G. Malnati, "MusA: Using indoor positioning and navigation to enhance cultural experiences in a museum," *Sensors*, vol. 2013, no. 13, pp. 17 445–17 471, 2013.

[13] C.-S. Wang, "An AR mobile navigation system integrating indoor positioning and content recommendation services," *World Wide Web*, vol. 22, pp. 1241–1262, 2019.

[14] E. Díaz, M. Pérez, D. Gualda, J. Villadangos, J. Ureña, and J. García, "Ultrasonic indoor positioning for smart environments: a mobile application," in *Proceeding of the Experiment@ International Conference*, Faro, Portugal, 2017, pp. 280–285.

[15] A. Cidronali, G. Collodi, M. Lucarelli, S. Maddio, M. Passafiume, and G. Pelosi, "Assessment of anchors constellation features in RSSI-based indoor positioning systems for smart environments," *Electronics*, vol. 9, no. ID 1026, pp. 1–18, 2020.

[16] C.T. Nguyen *et al.*, "A comprehensive survey of enabling and emerging technologies for social distancing - Part I: Fundamentals and enabling technologies," *IEEE Access*, vol. 8, pp. 153 479–153 507, 2020.

[17] ——, "A comprehensive survey of enabling and emerging technologies for social distancing - Part II: Emerging technologies and open issues," *IEEE Access*, vol. 8, pp. 154 209–154 236, 2020.

[18] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor localization using 802.11 time differences of arrival," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 614–623, 2015.

[19] S. He and S.-H. Gary Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[20] K.-M. Park, S.-H. Hyun, S. Lee, and S.-C. Kim, "Fine-resolution ranging scheme based on signal strength in indoor hallway with rough-surface slab waveguide," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 5502510, pp. 1–10, 2021.

[21] P. Kriz, F. Maly, and T. Kozel, "Improving indoor localization using bluetooth low energy beacons," *Mobile Information Systems*, vol. 2016, no. Article ID 2083094, pp. 1–11, 2016.

[22] G. S. de Blasio, A. Quesada-Arencibia, C. R. García, and J. C. Rodríguez-Rodríguez, "Bluetooth low energy technology applied to indoor positioning systems: An overview," in *Proceedings of the International Conference on Computer Aided Systems Theory*, Las Palmas de Gran Canaria, Spain, 2019, pp. 83–90.

[23] Y. You and C. Wu, "Hybrid indoor positioning system for pedestrians with swinging arms based on smartphone IMU and RSSI of BLE," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 9510615, pp. 1–15, 2021.

[24] Y. Zhou, C. L. Law, Y. L. Guan, and F. Chin, "Indoor elliptical localization based on asynchronous UWB range measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 1, pp. 248–257, 2011.

[25] A. Jiménez and F. Seco, "Comparing Ubisense, BeSpoon, and DecaWave UWB location systems: Indoor performance analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 2106–2117, 2017.

[26] F. Mazhar, M. G. Khan, and B. Sällberg, "Precise indoor positioning using UWB: A review of methods, algorithms and implementations," *Wireless Personal Communications*, vol. 97, pp. 4467–4491, 2017.

[27] A. F. C. Errington, B. L. F. Daku, and A. F. Prugger, "Initial position estimation using RFID tags: A least-squares approach," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 11, pp. 2863–2869, 2010.

[28] C.-H. Huang, L.-H. Lee, C. C. Ho, L.-L. Wu, and Z.-H. Lai, "Real-time RFID indoor positioning system based on Kalman-filter drift removal and Heron-bilateration location estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 4, pp. 728–739, 2015.

[29] G. Casella, B. Bigliardi, and E. Bottani, "The evolution of RFID technology in the logistics field: A review," *Procedia Computer Science*, vol. 200, pp. 1582–1592, 2022.

[30] M. Driusso, C. Marshall, M. Sabathy, F. Knutti, H. Mathis, and F. Babich, "Indoor positioning using LTE signals," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Alcalá de Henares, Spain, 2016, pp. 1–8.

[31] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018.

[32] Y. Boussad, M. N. Mahfoudi, A. Legout, L. Lizzi, F. Ferrero, and W. Dabbous, "Evaluating smartphone accuracy for RSSI measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 5501012, pp. 1–12, 2021.

[33] I. Fernández, M. Mazo, J. L. Lázaro, D. Pizarro, E. Santiso, P. Martín, and C. Losada, "Guidance of a mobile robot using an array of static cameras located in the environment," *Autonomous Robots*, vol. 23, no. 4, pp. 305–324, 2007.

[34] H. Deng, Q. Fu, Q. Quan, K. Yang, and K.-Y. Cai, "Indoor multi-camera-based testbed for 3-D tracking and control of UAVs," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3139–3156, 2020.

[35] A. Morar, A. Moldoveanu, I. Mocanu, F. Moldoveanu, I. E. Radoi, V. Asavei, A. Gradinaru, and A. Butean, "A comprehensive survey of indoor localization methods based on computer vision," *Sensors*, vol. 2020, no. Article ID 2641, pp. 1–36, 2020.

[36] E. Aparicio-Esteve, Á. Hernández, J. Ureña, and J. M. Villadangos, "Visible light positioning system based on a quadrant photodiode and encoding techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5589–5603, 2020.

[37] A. M. Rahman, T. Li, and Y. Wang, "Recent advances in indoor localization via visible lights: A survey," *Sensors*, vol. 2020, no. Article ID 1382, pp. 1–26, 2020.

[38] A. H. A. Bakar, T. Glass, H. Y. Tee, F. Alam, and M. Legg, "Accurate visible light positioning using multiple-photodiode receiver and machine learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 7500812, pp. 1–12, 2021.

[39] F. Höflinger, J. Müller, R. Zhang, L. M. Reindl, and W. Burgard, "A wireless micro inertial measurement unit (IMU)," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 9, pp. 2583–2595, 2013.

[40] A. Perttula, H. Leppäkoski, M. Kirkko-Jaakkola, P. Davidson, J. Collin, and J. Takala, "Distributed indoor positioning system with inertial measurements and map matching," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 11, pp. 2682–2695, 2014.

[41] X. Hou and J. Bergmann, "Pedestrian dead reckoning with wearable sensors: A systematic review," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 143–152, 2021.

[42] V. Pasku, A. De Angelis, G. De Angelis, D. D. Arumugam, M. Dionigi, P. Carbone, A. Moschitta, and D. S. Ricketts, "Magnetic field-based positioning systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 2003–2017, 2017.

[43] S.-C. Yeh, W.-H. Hsu, W.-Y. Lin, and Y.-F. Wu, "Study on an indoor positioning system using Earth's magnetic field," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 3, pp. 865–872, 2020.

[44] R. Kusche, S. O. Schmidt, and H. Hellbrück, "Indoor positioning via artificial magnetic fields," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 8502509, pp. 1–9, 2021.

[45] M. O. Khyam, L. Xinde, S. S. Ge, and M. R. Pickering, "Multiple access chirp-based ultrasonic positioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3126–3137, 2017.

[46] T. Aguilera, F. J. Álvarez, D. Gualda, J. M. Villadangos, Álvaro Hernández, and J. Ureña, "Multipath compensation algorithm for TDMA-based ultrasonic local positioning systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 5, pp. 984–991, 2018.

[47] M. Liu, L. Cheng, K. Qian, J. Wang, J. Wang, and Y. Liu, "Indoor acoustic localization: a survey," *Human-centric Computing and Information Sciences*, vol. 10, no. Article number 2, pp. 1–24, 2020.

[48] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1974–2002, 2017.

[49] G. Martín Mendoza-Silva, J. Torres-Sospedra, and J. Huerta, "A meta-review of indoor positioning systems," *Sensors-Basel*, vol. 19, no. ID 4507, pp. 1–45, 2019.

[50] F. Seco, J. C. Prieto, A. R. Jiménez, and J. Guevara, "Compensation of multiple access interference effects in CDMA-based acoustic positioning systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 10, pp. 2368–2378, 2014.

[51] D. Albuquerque, J. Vieira, S. Lopes, T. Aguilera, and F. Álvarez, "Doppler resilient modulation in a CDMA-based acoustic local positioning system," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Alcalá de Henares, Spain, 2016, pp. 1–8.

[52] J. Aparicio, T. Aguilera, and F. J. Álvarez, "Robust airborne ultrasonic positioning of moving targets in weak signal coverage areas," *IEEE Sensors Journal*, vol. 20, no. 21, pp. 13 119–13 130, 2020.

[53] R. B. Langley, "Dilution of precision," *GPS World*, vol. May, pp. 52–59, 1999.

[54] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.

[55] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 13–32, 2009.

[56] R. Mautz, "Indoor positioning technologies," Habilitation Thesis, available at https://doi.org/10.3929/ethz-a-007313554, 2012, last accessed on July 22, 2022.

[57] F. Ijaz, H. K. Yang, A. W. Ahmad, and C. Lee, "Indoor positioning: A review of indoor ultrasonic positioning systems," in *Proceedings of the International Conference on Advanced Communications Technology*, PyeongChang, South Korea, 2013, pp. 1146–1150.

[58] L. Mainetti, L. Patrono, and I. Sergi, "A survey on indoor positioning systems," in *Proceedings of the International Conference on Software, Telecommunications and Computer Networks*, Split, Croatia, 2014, pp. 111–120.

[59] J. Xiao, Z. Zhou, Y. Yi, and L. M. Ni, "A survey on wireless indoor localization from the device perspective," *ACM Computing Surveys*, vol. 49, no. ID 25, pp. 1–31, 2016.

[60] K. Mannay, N. Benhadjyoussef, M. Machhout, and J. Ureña, "Location and positioning systems: Performance and comparison," in *Proceedings of the International Conference on Control Engineering and Information Technology*, Hammamet, Tunisia, 2016, pp. 1–6.

[61] R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J. Fangmeyer Jr., "Evolution of indoor positioning technologies: A survey," *Journal of Sensors*, no. Article ID 2630413, pp. 1–21, 2017.

[62] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017.

[63] L. Dorji and T. Horanont, "Comparative assessment of indoor positioning technologies, techniques, and algorithms," in *Proceedings of the International Joint Symposium on Artificial Intelligence and Natural Language Processing*, Pattaya, Thailand, 2018, pp. 1–6.

[64] G. Oguntala, R. Abd-Alhameed, S. Jones, J. Noras, M. Patwary, and J. Rodriguez, "Indoor location identification technologies for real-time IoT-based applications: An inclusive survey," *Computer Science Review*, vol. 30, pp. 55–80, 2018.

[65] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[66] M. Pérez, D. Gualda, J. de Vicente, J. Villadangos, and J. Ureña, "Review of UAV positioning in indoor environments and new proposal based on US measurements," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Pisa, Italy, 2019, pp. 1–8.

[67] J. Mier, A. Jaramillo-Alcázar, and J. J. Freire, "At a glance: Indoor positioning systems technologies and their applications areas," in *Proceedings of the International Conference on Information Technology & Systems*, Quito, Ecuador, 2019, pp. 483–493.

[68] M. García-Catalá, M. Rodríguez-Sánchez, and E. Martín-Barroso, "Survey of indoor location technologies and wayfinding systems for users with cognitive disabilities in emergencies," *Behaviour and Information Technology*, pp. 1–25, 2020.

[69] W. C. Simões, G. S. Machado, A. M. Sales, M. M. de Lucena, N. Jazdi, and V. F. de Lucena, Jr., "A review of technologies and techniques for indoor navigation systems for the visually impaired," *Sensors*, vol. 2020, no. Article ID 3935, pp. 1–35, 2020.

[70] A. Billa, I. Shayea, A. Alhammadi, Q. Abdullah, and M. Roslee, "An overview of indoor localization technologies: Toward IoT navigation services," in *Proceedings of the International Symposium on Telecommunication Technologies*, Shah Alam, Malaysia, 2020, pp. 76–81.

[71] H. Obeidat, W. Shuaieb, O. Obeidat, and R. Abd-Alhameed, "A review of indoor localization techniques and wireless technologies," *Wireless Personal Communications*, vol. 119, pp. 289–327, 2021.

[72] Y. Bai, L. Lu, J. Cheng, J. Liu, Y. Chen, and J. Yu, "Acoustic-based sensing and applications: A survey," *Computer Networks*, vol. 181, no. Article ID 107447, pp. 1–19, 2020.

[73] S. Argentieri, P. Danès, and P. Souères, "A survey on sound source localization in robotics: From binaural to array processing methods," *Computer Speech and Language*, vol. 34, no. 1, pp. 87–112, 2015.

[74] C. Rascon and I. Meza, "Localization of sound sources in robotics: A review," *Robotics and Autonomous Systems*, vol. 96, pp. 184–210, 2017.

[75] T. Chong, X. Tang, C. Leng, M. Yogeswaran, O. Ng, and Y. Chong, "Sensor technologies and Simultaneous Localization and Mapping (SLAM)," in *Proceedings of the IEEE International Symposium on Robotics and Intelligent Sensors*, Langkawi, Malaysia, 2015, pp. 174–179.

[76] M. Zaffar, S. Ehsan, R. Stolkin, and K. McDonald Maier, "Sensors, SLAM and long-term autonomy: A review," in *Proceedings of the NASA/ESA Conference on Adaptive Hardware and Systems*, Edinburgh, United Kingdom, 2018, pp. 285–290.

[77] R. K. Ing, N. Quieffin, S. Catheline, and M. Fink, "In solid localization of finger impacts using acoustic time-reversal process," *Applied Physics Letters*, vol. 87, no. 204104, pp. 1–3, 2005.

[78] K. Kalgaonkar and B. Raj, "One-handed gesture recognition using ultrasonic Doppler sonar," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, 2009, pp. 1889–1892.

[79] W. Wang, A. X. Liu, and K. Sun, "Device-free gesture tracking using acoustic signals," in *Proceedings of the Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2016, pp. 82–94.

[80] E. Foxlin, M. Harrington, and G. Pfeifer, "Constellation: A wide-range wireless motion-tracking system for augmented reality and virtual set applications," in *Proceeding of the Annual Conference on Computer Graphics and Interactive Techniques*, Orlando, FL, USA, 1998, pp. 371–378.

[81] T. Sato, S. Nakamura, K. Terabayashi, M. Sugimoto, and H. Hashizume, "Design and implementation of a robust and real-time ultrasonic motion-capture system," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Guimarães, Portugal, 2011, pp. 1–4.

[82] H. Everett, D. DeMuth, and E. Stitz, "Survey of collision avoidance and ranging sensors for mobile robots, revision 1," Naval Command, Control and Ocean Surveillance Center, Technical Report 1194, 1992.

[83] W. E. Moritz, P. L. Shreve, and L. E. Mace, "Analysis of an ultrasonic spatial locating system," *IEEE Transactions on Instrumentation and Measurement*, vol. 25, no. 1, pp. 43–50, 1976.

[84] B. Everett, "A multi-element ultrasonic ranging array," Department of the Navy, Naval Sea Systems Command, Tech. Rep. NAVSEA No. 450-90G-TR-0001, 1985.

[85] F. Figueroa and J. S. Lamancusa, "A method for accurate detection of time of arrival: Analysis and design of an ultrasonic ranging system," *The Journal of the Acoustical Society of America*, vol. 91, no. 1, pp. 486–494, 1992.

[86] ——, "An ultrasonic ranging system for robot position sensing," in *Proceedings of the Meeting of the Acoustical Society of America*, Anaheim, CA, USA, 1986, p. S100.

[87] D. E. Manolakis, "Efficient solution and performance analysis of 3-D position estimation by trilateration," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1239–1248, 1996.

[88] F. Thomas and L. Ros, "Revisiting trilateration for robot localization," *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 93–101, 2005.

[89] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu, "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," *IEEE Communications Surveys and Tutorials*, vol. 22, no. 4, pp. 2634–2657, 2020.

[90] W. H. Foy, "Position-location solutions by Taylor-series estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-12, no. 2, pp. 187–194, 1976.

[91] K. Cheung, H. So, W. Ma, and Y. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1121–1128, 2004.

[92] R. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[93] H. R. Everett, "A microprocessor controlled autonomous sentry robot," M.Sc. thesis, Naval Postgraduate School, 1982.

[94] G. Giralt, R. Chatila, and M. Vaisset, "An integrated navigation and motion control system for autonomous multisensory mobile robots," in *Autonomous Robot Vehicles*. New York, NY, USA: Springer, 1990, pp. 420–443.

[95] J. M. Evans, "Helpmate: an autonomous mobile robot courier for hospitals," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Munich, Germany, 1994, pp. 1695–1700.

[96] D. Miller, "Two dimensional mobile robot positioning using onboard sonar," in *Proceedings of the 9th William T. Pecora Memorial Remote Sensing Symposium*, Sioux Falls, SD, USA, 1984, pp. 362–369.

[97] H. P. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Proceedings of the IEEE International Conference on Robotics and Automation*, St. Louis, MO, USA, 1985, pp. 116–121.

[98] J. L. Crowley, "Dynamic world modeling for an intelligent mobile robot using a rotating ultra-sonic ranging device," in *Proceeding of the IEEE International Conference on Robotics and Automation*, St. Louis, MO, USA, 1985, pp. 128–135.

[99] S. A. Walter, "The sonar ring: obstacle detection for a mobile robot," in *IEEE International Conference on Robotics and Automation*, Raleigh, NC, USA, 1987, pp. 1574–1579.

[100] CCS Robotics, "Speciminder," Available online at http://www.ccsrobotics.com/products/speciminder.html, 2008, last accessed on July 22, 2022.

[101] M. Drumheller, "Mobile robot localization using sonar," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 2, pp. 325–332, 1987.

[102] H. R. Beom and H. S. Cho, "Mobile robot localization using a single rotating sonar and two passive cylindrical beacons," *Robotica*, vol. 13, no. 3, pp. 243–252, 1995.

[103] J. A. Janét, D. S. Schudel, M. W. White, A. G. England, R. C. Luo, and W. E. Snyder, "Global self-localization for actual mobile robots: Generating and sharing topographical knowledge using the region-feature neural network," in *Proceedings of the IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Washington, DC, USA, 1996, pp. 619–626.

[104] H. Liu, F. Sun, B. Fang, and X. Zhang, "Robotic room-level localization using multiple sets of sonar measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 1, pp. 2–13, 2017.

[105] E. Dijk, K. van Berkel, R. Aarts, and E. van Loenen, "Estimation of 3D device position by analyzing ultrasonic reflection signals," in *Proceedings of the Annual Workshop on Circuits, Systems and Signal Processing*, Veldhoven, Netherlands, 2003, pp. 1–7.

[106] M. Rossi, J. Seiter, O. Amft, S. Buchmeier, and G. Tröster, "RoomSense: An indoor positioning system for smartphones using active sound probing," in *Proceedings of the ACM Augmented Human International Conference*, Stuttgart, Germany, 2013, pp. 89–95.

[107] T. Arai and E. Nakano, "Development of measuring equipment for location and direction (MELODI) using ultrasonic waves," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 105, no. 3, pp. 152–156, 1983.

[108] L. Kleeman, "Ultrasonic autonomous robot localization system," in *Proceedings of the IEEE/RSJ International Workshop on Intelligent Robots and Systems*, Tsukuba, Japan, 1989, pp. 212–219.

[109] ——, "Optimal estimation of position and heading for mobile robots using ultrasonic beacons and dead-reckoning," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Nice, France, 1992, pp. 2582–2587.

[110] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket location-support system," in *Proceedings of the International Conference on Mobile Computing and Networking*, Boston, MA, United States, 2000, pp. 32–43.

[111] A. Smith, H. Balakrishnan, M. Goraczko, and N. Priyantha, "Tracking moving devices with the Cricket location system," in *Proceeding of the International Conference on Mobile Systems, Applications, and Services*, Boston, MA, United States, 2004, pp. 190–202.

[112] C. Randell and H. L. Muller, "Low cost indoor positioning system," in *Proceedings of the International Conference on Ubiquitous Computing*, Atlanta, GA, United States, 2001, pp. 42–48.

[113] M. R. McCarthy and H. L. Muller, "RF free ultrasonic positioning," in *Proceedings of the IEEE International Symposium on Wearable Computers*, White Plains, NY, USA, 2003, pp. 79–85.

[114] G. Borriello, A. Liu, T. Offer, C. Palistrant, and R. Sharp, "WALRUS: Wireless acoustic location with room-level resolution using ultrasound," in *Proceedings of the International Conference on Mobile Systems, Applications, and Services*, Seattle, WA, USA, 2005, pp. 191–203.

[115] C. V. Lopes, A. Haghighat, A. Mandal, T. Givargis, and P. Baldi, "Localization of off-the-shelf mobile devices using audible sound: architectures, protocols and performance assessment," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 10, no. 2, pp. 38–50, 2006.

[116] Jan D. Bjerknes, Wenguo Liu, Alan FT Winfield and Chris Melhuish, "Low cost ultrasonic positioning system for mobile robots," in *Proceedings of Towards Autonomous Robotic Systems*, Aberystwyth, United Kingdom, 2007, pp. 107–114.

[117] E. Mangas and A. Bilas, "FLASH: fine-grained localization in wireless sensor networks using acoustic sound transmissions and high precision clock synchronization," in *Proceeding of the IEEE International Conference on Distributed Computing Systems*, Montreal, Canada, 2009, pp. 289–298.

[118] G. Oberholzer, P. Sommer, and R. Wattenhofer, "SpiderBat: Augmenting wireless sensor networks with distance and angle information," in *Proceedings of the ACM/IEEE International Conference on Information Processing in Sensor Networks*, Chicago, IL, USA, 2011, pp. 211–222.

[119] A. Sanchez, A. de Castro, S. Elvira, G. Glez-de-Rivera, and J. Garrido, "Autonomous indoor ultrasonic positioning system based on a low-cost conditioning circuit," *Measurement*, vol. 45, pp. 276–283, 2012.

[120] T. Akiyama, M. Nakamura, M. Sugimoto, and H. Hashizume, "Smart phone localization method using dual-carrier acoustic waves," in *Proceeding of the International Conference on Indoor Positioning and Indoor Navigation*, Montbeliard-Belfort, France, 2013, pp. 1–9.

[121] T. Akiyama, M. Sugimoto, and H. Hashizume, "SyncSync: Time-of-arrival based localization method using light-synchronized acoustic waves for smartphones," in *Proceeding of the International Conference on Indoor Positioning and Indoor Navigation*, Banff, Canada, 2015, pp. 1–9.

[122] C. Medina, J. C. Segura, and Á. De la Torre, "Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy," *Sensors*, vol. 13, pp. 3501–3526, 2013.

[123] A. Ens, F. Höflinger, J. Wendeberg, J. Hoppe, R. Zhang, A. Bannoura, L. M. Reindl, and C. Schindelhauer, "Acoustic self-calibrating system for indoor smart phone tracking," *International Journal of Navigation and Observation*, vol. 2015, no. Article ID 694695, pp. 1–15, 2015.

[124] H. Yang, R. Zhang, J. Bordoy, F. Höflinger, W. Li, C. Schindelhauer, and L. Reindl, "Smartphone-based indoor localization system using inertial sensor and acoustic transmitter/receiver," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 8051–8061, 2016.

[125] P. Lazik, N. Rajagopal, O. Shih, B. Sinopoli, and A. Rowe, "ALPS: a bluetooth and ultrasound platform for mapping and localization," in *Proceedings of the ACM Conference on Embedded Networked Sensor Systems*, Seoul, South Korea, 2015, pp. 73–84.

[126] J. Qi and G.-P. Liu, "A robust high-accuracy ultrasound indoor positioning system based on a wireless sensor network," *Sensors*, vol. 17, no. 2554, pp. 1–17, 2017.

[127] L. Kleeman and R. Kuc, "Sonar sensing," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Berlin, Germany: Springer, 2008, pp. 491–519.

[128] S. A. Hovanessian, *Radar system design and analysis*. Artech House, 1984.

[129] A. Ward, A. Jones, and A. Hopper, "A new location technique for the active office," *IEEE Personal Communications*, vol. 4, no. 5, pp. 42–47, 1997.

[130] J. Martín Abreu, R. Ceres, L. Calderón, M. Jiménez, and P. González-de-Santos, "Measuring the 3D-position of a walking vehicle using ultrasonic and electromagnetic waves," *Sensors and Actuators, A: Physical*, vol. 75, no. 2, pp. 131–138, 1999.

[131] A. R. Jiménez and F. Seco, "Precise localisation of archaeological findings with a new ultrasonic 3D positioning sensor," *Sensors and Actuators A: Physical*, vol. 123-124, pp. 224–233, 2005.

[132] A. Marco, R. Casas, J. Falcó, H. Gracia, J. I. Artigas, and A. Roy, "Location-based services for elderly and disabled people," *Computer Communications*, vol. 31, no. 6, pp. 1055–1066, 2008.

[133] R. Casas, A. Marco, J. J. Guerrero, and J. Falcó, "Robust estimator for non-line-of-sight error mitigation in indoor localization," *EURASIP Journal on Applied Signal Processing*, vol. 2006, no. Article ID 43429, pp. 1–8, 2006.

[134] J. Yang, S. Sidhom, G. Chandrasekaran, T. Vu, H. Liu, N. Cecan, Y. Chen, M. Gruteser, and R. P. Martin, "Detecting driver phone use leveraging car speakers," in *Proceedings of the Annual International Conference on Mobile Computing and Networking*, Las Vegas, NV, USA, 2011, pp. 97–108.

[135] B. Xu, R. Yu, G. Sun, and Z. Yang, "Whistle: Synchronization-free TDOA for localization," in *Proceedings of the IEEE International Conference on Distributed Computing Systems*, Minneapolis, MN, USA, 2011, pp. 760–769.

[136] M. Syafrudin, C. Walter, and H. Schweinzer, "Robust locating using LPS LOSNUS under NLOS conditions," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Busan, South Korea, 2014, pp. 582–590.

[137] S. I. Lopes, J. M. Vieira, J. Reis, D. Albuquerque, and N. B. Carvalho, "Accurate smartphone indoor positioning using a WSN infrastructure and non-invasive audio for TDoA estimation," *Pervasive and Mobile Computing*, vol. 20, pp. 29–46, 2015.

[138] A. De Angelis, A. Moschitta, P. Carbone, M. Calderini, S. Neri, R. Borgna, and M. Peppucci, "Design and characterization of a portable ultrasonic indoor 3-D positioning system," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 10, pp. 2616–2625, 2015.

[139] H. Han, S. Yi, Q. Li, G. Shen, Y. Liu, and E. Novak, "AMIL: Localizing neighboring mobile devices through a simple gesture," in *Proceedings*

of the IEEE International Conference on Computer Communications, San Francisco, CA, USA, 2016, pp. 1–9.

[140] Y. Wang, J. Li, R. Zheng, and D. Zhao, "ARABIS: An asynchronous acoustic indoor positioning system for mobile devices," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Sapporo, Japan, 2017, pp. 1–8.

[141] C. Cai, R. Zheng, J. Li, L. Zhu, H. Pu, and M. Hu, "Asynchronous acoustic localization and tracking for mobile targets," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 830–845, 2020.

[142] G. Li, L. Zhang, F. Lin, M. Chen, and Z. Wang, "SAILoc: A novel acoustic single array system for indoor localization," in *Proceedings of the International Conference on Wireless Communications and Signal Processing*, Nanjing, China, 2017, pp. 1–6.

[143] R. Carotenuto, M. Merenda, D. Iero, and F. G. Della Corte, "An indoor ultrasonic system for autonomous 3-D positioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 7, pp. 2507–2518, 2019.

[144] X. Chen, Y. Chen, S. Cao, L. Zhang, X. Zhang, and X. Chen, "Acoustic indoor localization system integrating TDMA+FDMA transmission scheme and positioning correction technique," *Sensors*, vol. 19, no. Article ID 2353, pp. 1–18, 2019.

[145] H. Wang, L. Zhang, Z. Wang, and X. Luo, "PALS: High-accuracy pedestrian localization with fusion of smartphone acoustics and PDR," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Pisa, Italy, 2019, pp. 1–8.

[146] L. Zhang, M. Chen, X. Wang, and Z. Wang, "TOA estimation of chirp signal in dense multipath environment for low-cost acoustic ranging," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 2, pp. 355–367, 2019.

[147] S. Ogiso, K. Mizutani, N. Wakatsuki, and T. Ebihara, "Robust indoor localization in a reverberant environment using microphone pairs and asynchronous acoustic beacons," *IEEE Access*, vol. 7, pp. 123 116–123 127, 2019.

[148] R. Chen, Z. Li, F. Ye, G. Guo, S. Xu, L. Qian, Z. Liu, and L. Huang, "Precise indoor positioning based on acoustic ranging in smartphone," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. Article ID 9509512, pp. 1–12, 2021.

[149] H. Murakami, T. Suzaki, M. Nakamura, H. Hashizume, and M. Sugimoto, "Five degrees-of-freedom pose-estimation method for smartphones using a single acoustic anchor," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8030–8044, 2021.

[150] K. Audenaert, H. Peremans, Y. Kawahara, and J. M. Van Campenhout, "Accurate ranging of multiple objects using ultrasonic sensors," in *Proceeding of the IEEE International Conference on Robotics and Automation*, Nice, France, 1992, pp. 1733–1738.

[151] H. Peremans, K. Audenaert, and J. M. Van Campenhout, "A high-resolution sensor based on tri-aural perception," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 1, pp. 36–48, 1993.

[152] J. Ureña, M. Mazo, J. J. García, A. Hernández, and E. Bueno, "Classification of reflectors with an ultrasonic sensor for mobile robot applications," *Robotics and Autonomous Systems*, vol. 29, no. 4, pp. 269–279, 1999.

[153] A. Jiménez, A. Hernández, J. Ureña, M. C. Pérez, F. J. Álvarez, C. de Marziani, J. J. García, and J. M. Villadangos, "EMFi-based ultrasonic transducer for robotics applications," *Sensors and Actuators, A: Physical*, vol. 148, pp. 342–349, 2008.

[154] F. J. Álvarez, J. Ureña, J. J. García, M. Mazo, C. De Marziani, A. Herández, and J. M. Villadangos, "A comparative analysis of two modulation schemes for the efficient transmission of complementary sequences in a pulse compression ultrasonic system," in *Proceeding of the International Conference on Telecommunications and Computer Networks*, San Sebastián, Spain, 2004, pp. 1–5.

[155] K.-W. Jörg, M. Berg, and M. Müller, "Using pseudo-random codes for mobile robot sonar sensing," in *Proceeding of the IFAC Symposium on Intelligent Autonomous Vehicles*, Madrid, Spain, 1998, pp. 1–6.

[156] A. Hernández, J. Ureña, J. J. García, M. Mazo, D. Hernanz, J.-P. Dérutin, and J. Sérot, "Ultrasonic ranging sensor using simultaneous emissions from different transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 51, no. 12, pp. 1660–1670, 2004.

[157] J. Lahtonen, "On the odd and the aperiodic correlation properties of the kasami sequences," *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1506–1508, 1995.

[158] B. M. Popović, "Efficient Golay correlator," *IEEE Electronic Letters*, vol. 35, no. 17, pp. 1427–1428, 1999.

[159] F. J. Álvarez, J. Ureña, M. Mazo, Á. Herández, J. J. García, and J. A. Jiménez, "Efficient generator and pulse compressor for complementary sets of four sequences," *IEEE Electronic Letters*, vol. 40, no. 11, pp. 703–704, 2004.

[160] C. De Marziani, J. Ureña, Á. Hernández, M. Mazo, F. J. Álvarez, J. J. García, and P. Donato, "Modular architecture for efficient generation and correlation of complementary set of sequences," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2323–2337, 2007.

[161] M. C. Pérez, J. Ureña, A. Hernández, C. De Marziani, J. J. García, and A. Jiménez, "Optimized correlator for LS codes-based CDMA systems," *IEEE Communications Letters*, vol. 15, no. 2, pp. 223–225, 2011.

[162] M. C. Pérez, J. Ureña, A. Hernández, A. Jiménez, D. Ruiz, F. J. Álvarez, and C. De Marziani, "Performance comparison of different codes in an ultrasonic positioning system using using DS-CDMA," in *Proceeding of the IEEE International Symposium on Intelligent Signal Processing*, Budapest, Hungary, 2009, pp. 125–130.

[163] C. Zhang, X. Lin, S. Yamada, and M. Hatori, "General method to construct LS codes by complete complementary sequences," *IEICE Transactions on Communications*, vol. E88-B, no. 8, pp. 3484–3487, 2005.

[164] S. Murano, C. Pérez-Rubio, D. Gualda, F. J. Álvarez, T. Aguilera, and C. De Marziani, "Evaluation of Zadoff–Chu, Kasami, and chirp-based encoding schemes for acoustic local positioning systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5356–5368, 2020.

[165] T. O'Farrell, "Design and evaluation of a high data rate optical wireless system for the diffuse indoor channel using barker spreading codes and RAKE reception," *IET Communications*, vol. 2, no. 1, pp. 35–44, 2008.

[166] D. A. Shaddock, "Digitally enhanced heterodyne interferometry," *Optics Letters*, vol. 32, no. 22, pp. 3355–3357, 2007.

[167] G. de Vine, D. S. Rabeling, B. J. Slagmolen, T. T. Lam, S. Chua, D. M. Wuchenich, D. E. McClelland, and D. A. Shaddock, "Picometer level displacement metrology with digitally enhanced heterodyne interferometry," *Optics Express*, vol. 17, no. 2, pp. 828–837, 2009.

[168] A. Sutton, K. McKenzie, B. Ware, and D. A. Shaddock, "Laser ranging and communications for LISA," *Optics Express*, vol. 18, no. 20, pp. 20 759–20 773, 2010.

[169] E. García, J. J. García, J. Ureña, M. C. Pérez, and D. Ruiz, "Multilevel complementary sets of sequences and their application in UWB," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Zürich, Switzerland, 2010, pp. 1–5.

[170] M. Hazas and A. Ward, "A novel broadband ultrasonic location system," in *Proceeding of the International Conference on Ubiquitous Computing*, Gothenburg, Sweden, 2002, pp. 264–280.

[171] M. Hazas and A. Hopper, "Broadband ultrasonic location systems for improved indoor positioning," *IEEE Transactions on Mobile Computing*, vol. 5, no. 5, pp. 536–547, 2006.

[172] O. A. M. Aly and A. S. Omar, "Spread spectrum ultrasonic positioning system," in *Proceeding of the Workshop on Positioning, Navigation and Communication*, Hannover, Germany, 2005, pp. 109–114.

[173] E. O. Dijk, C.H. van Berkel, R. M. Aarts, and E. J. van Loenen, "A 3-D indoor positioning method using a single compact base station," in *Proceedings of the IEEE Annual Conference on Pervasive Computing and Communications*, Orlando, FL, USA, 2004, pp. 101–110.

[174] J. C. Prieto, A. R. Jiménez, J. Guevara, J. L. Ealo, F. Seco, J. O. Roa, and F. Ramos, "Performance evaluation of 3D-LOCUS advanced acoustic LPS," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 8, pp. 2385–2395, 2009.

[175] J. Gonzalez and C. Bleakley, "High-precision robust broadband ultrasonic location and orientation estimation," *IEEE Journal on Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 832–844, 2009.

[176] M. M. Saad, C. J. Bleakley, and S. Dobson, "Robust high-accuracy ultrasonic range measurement system," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 10, pp. 3334–3341, 2011.

[177] S. J. Kim and B. K. Kim, "Accurate hybrid global self-localization algorithm for indoor mobile robots with two-dimensional isotropic ultrasonic receivers," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 10, pp. 3391–3404, 2011.

[178] ——, "Dynamic ultrasonic hybrid localization system for indoor mobile robots," *IEEE Transactions on Industrial Electronicst*, vol. 60, no. 10, pp. 4562–4573, 2013.

[179] P. Lazik and A. Rowe, "Indoor pseudo-ranging of mobile devices using ultrasonic chirps," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, Toronto, Canada, 2012, pp. 99–112.

[180] I. Rishabh, D. Kimber, and J. Adcock, "Indoor localization using controlled ambient sounds," in *Proceedings of the International Conference*

on Indoor Positioning and Indoor Navigation, Sydney, Australia, 2012, pp. 1–10.

[181] C. Sertatıl, M. A. Altınkaya, and K. Raoof, "A novel acoustic indoor localization system employing CDMA," *Digital Signal Processing*, vol. 22, no. 3, pp. 506–517, 2012.

[182] S. Widodo, T. Shiigi, N. Hayashi, H. Kikuchi, K. Yanagida, Y. Nakatsuchi, Y. Ogawa, and N. Kondo, "Moving object localization using sound-based positioning system with doppler shift compensation," *MDPI Robotics*, vol. 2, no. 2, pp. 36–53, 2013.

[183] Z. Huang, L. W. J. Tsay, T. Shiigi, X. Zhao, H. Nakanishi, T. Suzuki, Y. Ogawa, and N. Kondo, "A noise tolerant spread spectrum sound-based local positioning system for operating a quadcopter in a greenhouse," *Sensors*, vol. 20, no. ID 1981, pp. 1–15, 2020.

[184] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling fine-grained smartphone localization via acoustic anchors," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1144–1156, 2016.

[185] J. Ureña, A. Hernández, A. Jiménez, J. Villadangos, M. Mazo, J. García, J. García, F. Álvarez, C. De Marziani, M. Pérez, J. Jiménez, A. Jiménez, and F. Seco, "Advanced sensorial system for an acoustic LPS," *Microprocessors and Microsystems*, vol. 31, pp. 393–401, 2007.

[186] D. Ruiz, E. García, J. Ureña, D. de Diego, D. Gualda, and J. García, "Extensive ultrasonic local positioning system for navigating with mobile robots," in *Proceedings of the IEEE Workshop on Positioning, Navigation and Communication*, Dresden, Germany, 2013, pp. 1–6.

[187] J. Ureña, A. Hernández, J. J. García, J. M. Villadangos, M. Pérez, D. Gualda, F. J. Álvarez, and T. Aguilera, "Acoustic local positioning with encoded emission beacons," *Proceedings of the IEEE*, vol. 106, no. 6, pp. 1042–1062, 2018.

[188] D. Gualda, M. C. Pérez-Rubio, J. Ureña, S. Pérez-Bachiller, J. M. Villadangos, Á. Hernández, J. J. García, and A. Jiménez, "LOCATE-US: Indoor positioning for mobile devices using encoded ultrasonic signals, inertial sensors and graph-matching," *Sensors*, vol. 21, no. ID 1950, pp. 1–24, 2021.

[189] J. Moutinho, R. Araújo, and D. Freitas, "Indoor localization with audible sound – towards practical implementation," *Pervasive and Mobile Computing*, vol. 29, pp. 1–16, 2016.

[190] T. Aguilera, F. Seco, F. J. Álvarez, and A. Jiménez, "Broadband acoustic local positioning system for mobile devices with multiple access interference cancellation," *Measurement*, vol. 116, pp. 483–494, 2018.

[191] T. Aguilera, J. Aparicio, and F. J. Álvarez, "High availability acoustic positioning for targets moving in poor coverage areas," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, ID 9700511, pp. 1–11, 2021.

[192] D. Esslinger, P. Rapp, S. Wiertz, H. Rendich, R. Marsden, O. Sawodny, and C. Tarín, "Accurate optoacoustic and inertial 3-D pose tracking of moving objects with particle filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 3, pp. 893–906, 2020.

[193] S. Holm, "Airborne ultrasound data communications: The core of an indoor positioning system," in *Proceedings of the IEEE International Ultrasonics Symposium*, Rotterdam, Netherlands, 2005, pp. 1–4.

[194] C. Li, D. A. Hutchins, and R. J. Green, "Short-range ultrasonic digital communications in air," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, no. 4, pp. 908–918, 2008.

[195] ——, "Short-range ultrasonic communications in air using quadrature modulation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 10, pp. 2060–2072, 2009.

[196] D. W. Wax, "MFSK - The basis for robust acoustical communications," in *Proceeding of the IEEE/MTS OCEANS Conference*, Boston, MA, United States, 1981, pp. 61–66.

[197] K. F. Scussel, J. A. Rice, and S. Merriam, "A new MFSK acoustic modem for operation in adverse underwater channels," in *Proceedings of the IEEE/MTS OCEANS Conference*, Halifax, NS, Canada, 1997, pp. 247–254.

[198] S. Holm, "Ultrasound positioning based on time-of-flight and signal strength," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Sydney, Australia, 2012, pp. 1–6.

[199] S. Holm, O. B. Hovind, S. Rostad, and R. Holm, "Indoors data communications using airborne ultrasound," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Philadelphia, PA, USA, 2005, pp. 1–4.

[200] S. Holm, "Hybrid ultrasound-RFID indoor positioning: Combining the best of both worlds," in *Proceedings of the IEEE International Conference on RFID*, Orlando, FL, USA, 2009, pp. 155–162.

[201] A. Borrelli, C. Monti, M. Vari, and F. Mazzenga, "Channel models for IEEE 802.11b indoor system design," in *Proceedings of the IEEE International Conference of Communications*, Paris, France, 2004, pp. 3701–3705.

[202] E. Elnahrawy, X. Li, and R. P. Martin, "The limits of localization using signal strength: A comparative study," in *Proceeding of the IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, Santa Clara, CA, USA, 2004, pp. 406–414.

[203] P. Davidson and R. Piché, "A survey of selected indoor positioning methods for smartphones," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1347–1370, 2017.

[204] A. Hammoud, M. Deriaz, and D. Konstantas, "Robust ultrasound-based room-level localization system using COTS components," in *Proceedings of the International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services*, Shanghai, China, 2016, pp. 11–19.

[205] S. Holm and C.-I. C. Nilsen, "Robust ultrasonic indoor positioning using transmitter arrays," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Zürich, Switzerland, 2010, pp. 1–5.

[206] M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama, "DOLPHIN: a practical approach for implementing a fully distributed indoor ultrasonic positioning system," in *Proceedings of the International Conference on Ubiquitous Computing*, Nottingham, United Kingdom, 2004, pp. 347–365.

[207] C. Medina, J. Segura, and S. Holm, "Feasibility of ultrasound positioning based on signal strength," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Sydney, Australia, 2012, pp. 1–9.

[208] S. P. Tarzia, P. A. Dinda, R. P. Dick, and G. Memik, "Indoor localization without infrastructure using the Acoustic Background Spectrum," in *Proceeding of the International Conference on Mobile Systems, Applications, and Services*, Bethesda, MD, United States, 2011, pp. 155–168.

[209] A. H. Moore, M. Brookes, and P. A. Naylor, "Roomprints for forensic audio applications," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, United States, 2013, pp. 1–4.

[210] Y.-C. Tung and K. G. Shin, "EchoTag: Accurate Infrastructure-Free Indoor Location Tagging with Smartphones," in *Proceedings of the Annual International Conference on Mobile Computing and Networking*, Paris, France, 2015, pp. 525–536.

[211] L. Phillips, C. B. Porter, N. Kottege, M. D'Souza, and M. Ros, "Machine learning based acoustic sensing for indoor room localisation using mobile phones," in *Proceedings of the International Conference on Sensing Technology*, Auckland, New Zealand, 2015, pp. 456–460.

[212] K. I. Gettapola, R. R. W. M. H. D. Ranaweera, and G. M. R. I. Godaliyadda, "Location based fingerprinting techniques for indoor positioning," in *Proceedings of the National Conference on Technology and Management*, Malabe, Sri Lanka, 2017, pp. 175–180.

[213] R. Leonardo, M. Barandas, and H. Gamboa, "A framework for infrastructure-free indoor localization based on pervasive sound analysis," *IEEE Sensors Journal*, vol. 18, no. 10, pp. 4136–4144, 2018.

[214] Y. Nagama, T. Umezawa, and N. Osawa, "Indoor localization based on analysis of environmental ultrasound," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Pisa, Italy, 2019, pp. 1–8.

[215] M. Jiang, J. Lundgren, S. Pasha, M. Carratù, C. Liguori, and G. Thungström, "Indoor silent object localization using ambient acoustic noise fingerprinting," in *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference*, Dubrovnik, Croatia, 2020, pp. 1–6.

[216] M. Nakamura, K. Fujimoto, H. Murakami, H. Hashizume, and M. Sugimoto, "Indoor localization method for a microphone using a single speaker," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation*, Lloret de Mar, Spain, 2021, pp. 1–8.

[217] K. Kunze and P. Lukowicz, "Symbolic object localization through active sampling of acceleration and sound signatures," in *Proceedings of the International Conference on Ubiquitous Computing*, Innsbruck, Austria, 2007, pp. 163–180.

[218] W. Huang, Y. Xiong, X.-Y. Li, H. Lin, X. Mao, P. Yang, Y. Liu, and X. Wang, "Swadloon: Direction finding and indoor localization using acoustic signal by shaking smartphones," *IEEE Transactions on Mobile Computing*, vol. 14, no. 10, pp. 2145–2157, 2015.

[219] Savi Technology, Inc., "Method and apparatus for locating items," patent number 5,528,232, filled on May 20th, 1994, granted on Jun. 18th, 1996.

[220] Netmor Ltd., "Ultrasonic positioning and tracking system," patent number 6,141,293, filled on Oct. 29th, 1998, granted on Oct. 31st, 2000.

[221] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper, "Implementing a sentient computing system," *Computer*, vol. 34, no. 8, pp. 50–56, 2001.

[222] Filetrac AS, "System for supervision and control of objects or persons," patent number US 6,433,689 B1, filled on Apr. 16th, 1999, granted on Aug. 13th, 2002.

[223] D. Clarke and A. Park, "Active-RFID system accuracy and its implications for clinical applications," in *Proceedings of IEEE Symposium on Computer-Based Medical Systems*, Salt Lake City, UT, USA, 2006, pp. 21–26.

[224] L. Greenemeier, "A positioning system that goes where GPS can't," Scientific American, available online at https://www.scientificamerican.com/article/indoor-positioning-system/, 2008, last accessed on July 22, 2022.

[225] Sonitor Technologies AS, "Sonitor Sense RTLS: Technology brief," 2014, available online at https://www.sonitor.com/s/100820-Sonitor-Sonitor-Sense-Technology-Brief-6s93.pdf. Last accessed on July 22, 2022.

[226] C. Brignone, T. Connors, G. Lyon, and S. Pradhan, "SmartLOCUS: An autonomous, self-assembling sensor network for indoor asset and systems management," Hewlett-Packard Laboratories, Tech. Rep. HPL-2003-41, 2005.

[227] J. Zhao and Y. Wang, "Autonomous ultrasonic indoor tracking system," in *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing with Applications*, Sydney, Australia, 2008, pp. 532–539.

[228] Celltek Electronics PTY Limited, "A tracking system," international publication number WO 2009/029980 A1, filled on Sept. 2nd, 2008, granted on Mar. 12nd, 2009.

[229] Aeroscout Ltd., "Improved room separation in a WLAN based RTLS and method therefore," European patent application EP 2151696 A1, 2010.

[230] Stanley Healthcare, "AeroScout Asset Management," Available online at http://www.stanleyhealthcare.com/resources/aeroscout-asset-management-solution-overview, last accessed on July 22, 2022.

[231] ——, "Stanley Healthcare and Sonitor partner to transform RTLS technology," Available online at https://www.stanleyhealthcare.com/stanley-healthcare-sonitor-expand-partnership, 2018, last accessed on July 22, 2022.

[232] Intelligent Sciences Ltd., "Ultrasonic in-building positioning system based on phase difference array with ranging," patent number US 7,796,471 B2, filled on Feb. 5th, 2009, granted on Sep. 14th, 2010.

[233] KT Corporation, "Positioning system using ultrasonic waves and method for operating the same," patent number US 7,764,574 B2, filled on Jul. 11th, 2008, granted on Jul. 27th, 2010.

[234] Centrak, Inc., "Methods and systems for synchronized ultrasonic real time location," patent number US 8,604,909 B1, filled on Jan. 7th, 2011, granted on Dec. 10th, 2013.

[235] CenTrak, Inc., "Hybrid IR-US RTLS System," patent number US 10,794,987 B2, filled on Feb. 13th, 2018, granted on Oct. 6th, 2020.

[236] Marvelmind Robotics, "Indoor "GPS"," Available online at https://marvelmind.com/pics/marvelmind_presentation.pdf, 2014, last accessed on July 22, 2022.

[237] D. Kang and Y.-J. Cha, "Autonomous UAVs for Structural Health Monitoring Using Deep Learning and an Ultrasonic Beacon System with Geo-Tagging," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 10, pp. 885–902, 2018.

[238] Hexamite, "Hx19 RFID Ultrasonic Positioning," 2016, available online at: https://www.hexamite.com/pdfs/hx19flyer.pdf, last accessed July 22, 2022.

[239] Certis Cisco Security Pte Ltd., "System and method for determining a location of a mobile device based on audio localization techniques," pub. number US 2020/0386843 A1, filled on Jun. 17th, 2018, granted on Dec. 10th, 2020.

[240] Sonitor Technologies AS, "Position determination system having a deconvolution decoder," patent number US 10,670,693 B2, filled on Dec. 29th, 2017, granted on Jun. 2nd, 2020.

[241] ——, "General method for Doppler correction of an acoustic signal based on deconvolution," patent pending.

[242] ——, "An optimal positioning engine for an acoustic ranging system," patent pending.

[243] V. Thio, J. Aparicio, K. B. Ånonsen, and J. K. Bekkeng, "Experimental evaluation of the Forkbeard indoor positioning system," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, ID 8500313, pp. 1–13, 2022.

[244] V. Thio, J. Aparicio, K. B. Ånonsen, J. K. Bekkeng, and W. Booij, "Fusing of a continuous output PDR algorithm with an ultrasonic positioning system," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2464–2474, 2022.

[245] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," *IEEE Computer Graphics and Applications*, vol. 22, no. 6, pp. 24–38, 2002.

[246] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Prentice Hall, 1993.

[247] S. Holm, *Waves with Power-Law Attenuation*. Springer, 2019.

[248] H. E. Bass and L. N. Bolen, "Ultrasonic background noise in industrial environments," *The Journal of the Acoustical Society of America*, vol. 78, no. 6, pp. 2013–2016, 1985.

[249] S. Holm, "Optimum FFT-based frequency acquisition with application to COSPAS-SARSAT," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 2, pp. 464–475, 1993.

[250] Murata, "Data Sheet MA40S4S/MA40S4R," Available online at https://www.murata.com/~/media/webrenewal/products/sensor/ultrasonic/open/datasheet_maopn.ashx?la=en, last accessed July 22, 2022.

[251] B. W. Lawton, "Damage to human hearing by airborne sound of very high frequency or ultrasonic frequency," Contract Research Report 343/2001, 2001.

[252] T. G. Leighton, B. Lineton, C. Dolder, and M. D. Fletcher, "Public exposure to airborne ultrasound and very high frequency sound," *Acoustics Today*, vol. 16, no. 3, pp. 17–26, 2020.

[253] E. García, S. Holm, J. J. García, and J. Ureña, "Link budget for low bandwidth and coded ultrasonic indoor location systems," in *Proceeding of the International Conference on Indoor Positioning and Indoor Navigation*, Guimarães, Portugal, 2011, pp. 1–4.

[254] M. C. Pérez-Rubio, C. Losada-Gutiérrez, F. Espinosa, J. Macias-Guarasa, J. Tiemann, F. Eckermann, C. Wietfeld, M. Katkov, S. Huba, J. Ureña, J. M. Villadangos, D. Gualda, E. Díaz, R. Nieto, E. Santiso, P. del Portillo, and M. Martínez, "A realistic evaluation of indoor robot position tracking systems: The IPIN 2016 competition experience," *Measurement*, vol. 135, pp. 151–162, 2019.

[255] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned," in *Proceeding of the International Conference on Information Processing in Sensor Networks*, Seattle, WA, United States, 2015, pp. 178–189.

[256] H. W. Wehn and P. R. Bélanger, "Ultrasound-based robot position estimation," *IEEE Transactions on Robotics and Automation*, vol. 13, no. 5, pp. 682–692, 1997.

[257] D. A. Bohn, "Environmental effects on the speed of sound," *Journal of the Audio Engineering Society*, vol. 36, no. 4, 1988.

[258] O. Cramer, "The variation of the specific heat ratio and the speed of sound in air with temperature, pressure, humidity, and CO_2 concentration," *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2510–2516, 1993.

[259] G. P. Howell and C. L. Morfey, "Frequency dependence of the speed of sound in air," *The Journal of the Acoustical Society of America*, vol. 82, no. 1, pp. 375–376, 1987.

[260] F. J. Álvarez and R. Kuc, "Dispersion relation for air via Kramers-Kronig analysis," *The Journal of the Acoustical Society of America*, vol. 124, no. 2, pp. EL57–EL61, 2008.

Joaquín Aparicio (M'17) received his Physics degree from the University of Extremadura, Spain, in 2008, and his Ph.D. degree in Electronics from the University of Alcalá, Spain, in 2014. From 2015 to 2018 he was a Postdoctoral Researcher at the Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan. Since 2018 he is a Postdoctoral Fellow at the Department of Informatics, University of Oslo, Oslo, Norway. His research interests include positioning systems, acoustic propagation, signal processing and communication systems.

Fernando J. Álvarez (M'07 - SM'17) received his Physics degree from the University of Sevilla, Spain, his Electronic Engineering degree from the University of Extremadura, Spain, and his Ph.D. degree in Electronics from the University of Alcalá, Spain. He also holds a M.Sc. degree in Signal Theory and Communications from the University of Vigo, Spain, and a M.Sc. degree in Telecommunications Engineering from the Universitat Oberta de Catalunya, Spain. Since 2001, he has been with the Department of Electrical Engineering, Electronics and Automation at the University of Extremadura, where he is currently a Full Professor and Head of the Sensory Systems Research Group. In 2008, he joined the Intelligent Sensors Laboratory, Yale University, New Haven, CT, USA, as a Post-Doctoral "José Castillejo" Fellow. His current research interests include local positioning systems, acoustic signal processing, and embedded computing.

Álvaro Hernández (Senior Member, IEEE) received the Ph.D. degree from the University of Alcalá, Spain, and from Blaise Pascal University, France, in 2003. He is currently a Professor of Digital Systems and Electronic Design with the Electronics Department, University of Alcalá. His research areas are ultrasonic sensory systems, electronic systems for mobile robots, and digital embedded systems.

Sverre Holm was born in Oslo, Norway, in 1954 and studied electrical engineering at the Norwegian Institute of Technology (NTNU), Trondheim and has a Ph. D. from 1982. He has academic experience from NTNU and Yarmouk University in Jordan (1984-86). Since 1995 he has been a professor of signal processing and acoustic imaging at the University of Oslo, first in the Informatics Department and lately in the Physics Department. In 2002 he was elected a member of the Norwegian Academy of Technological Sciences. He is a senior member of IEEE and has been an associate editor of IEEE Ultrasonics, Ferroelectric and Frequency Control. His industry experience includes GE Vingmed Ultrasound (1990-94) working on digital ultrasound imaging, and Sonitor Technologies (2000-05) where he developed ultrasonic indoor positioning. He is involved with several startups in the Oslo area working in acoustics and ultrasonics. His book "Waves with Power-Law Attenuation" was published in 2019 by Springer and the Acoustical Society of America. His research interests are medical imaging with ultrasound and elastography, modeling of waves in complex media, fractional calculus, and ultrasonic positioning.