Abstract—We present a novel sound localization algorithm for a non-line-of-sight (NLOS) sound source in indoor environments. Our approach exploits the diffraction properties of sound waves as they bend around a barrier or an obstacle in the scene. We combine a ray tracing based sound propagation algorithm with a Uniform Theory of Diffraction (UTD) model, which simulate bending effects by placing a virtual sound source on a wedge in the environment. We precompute the wedges of a reconstructed mesh of an indoor scene and use them to generate diffraction acoustic rays to localize the 3D position of the source. Our method identifies the convergence region of those generated acoustic rays as the estimated source position based on a particle filter. We have evaluated our algorithm in multiple scenarios consisting of a static and dynamic NLOS sound source. In our tested cases, our approach can localize a source position with an average accuracy error, 0.7m, measured by the L2 distance between estimated and actual source locations in a 7m×7m×3m room. Furthermore, we observe 37% to 130% improvement in accuracy over a state-of-the-art localization method that does not model diffraction effects, especially when a sound source is not visible to the robot.

I. INTRODUCTION

As mobile robots are increasingly used for different applications, there is considerable interest in developing new and improved methods for localization. The main goal is to compute the current location of the robot with respect to its environment. Localization is a fundamental capability required by an autonomous robot, as the current location is used to guide the future movement or actions. We assume that a map of the environment is given and different sensors on the robot are used to estimate its position and orientation in the environment. Some of the commonly used sensors include GPS, CCD or depth cameras, acoustics, etc. In particular, there is considerable work on using acoustic sensors for localization, including sonar signal processing for underwater localization and microphone arrays for indoor and outdoor scenes. In particular, the recent use of smart microphones in commodity or IoT devices (e.g., Amazon Alexa) has triggered interest in better acoustic localization methods [1], [2].

The acoustic sensors use the properties of sound waves to compute the source location. As the sound waves are emitted from a source, they transmit through the media and either reach the listener or microphone locations as direct paths, or after undergoing different wave effects including reflections, interference, diffraction, scattering, etc. Some of the earliest work on sound source localization (SSL) makes use of the time difference of arrival (TDOA) at the receiver [3], [4]. These methods only exploit the direct sound and its direction at the receiver, and do not take into account of reflections or other wave effects. As a result, it does not provide sufficient accuracy for many applications. Other techniques have been proposed to localize the position under different constraints or sensors [5], [6], [7], [8]. This includes modeling of higher order specular reflections [8] based on ray tracing and can model indirect sound effects.

In many scenarios, the sound source is not directly in line of sight of the listener (i.e. NLOS) and is occluded by obstacles. In such cases, there may not be much contribution in terms of direct sound, and simple methods based on

Diffraction-Aware Sound Localization for a Non-Line-of-Sight Source

Inkyu An1, Doheon Lee2, Jung-woo Choi3, Dinesh Manocha4, and Sung-eui Yoon5

1I. An, 2D. Lee, and 5S. Yoon (Corresponding author) are with the School of Computing, KAIST, Daejeon, South Korea; 3J. Choi is with the School of Electrical Engineering, KAIST; 4D. Manocha is with the Dept. of CS & ECE, Univ. of Maryland at College Park, USA: {inkyuan, doheonlee, jwoo}@kaist.ac.kr, dmanocha@gmail.com, sungeui@kaist.edu

Fig. 1. These figures show the testing environment (7m by 7m with 3m height) (a) and the accuracy error of our method with the dynamically moving sound source (b). The source moves along the red trajectory, and the obstacle causes the invisible area for the dynamic source. Invisibility of the source occurs from 27s to 48s, where our method maintains a high accuracy, while the prior method deteriorates due to the diffraction: the average distance errors of our and the prior method are 0.95m and 1.83m.
Sound source localization (SSL). Over the past two decades, many approaches have used time difference of arrival (TDOA) to localize sound sources. Knapp et al. presented a good estimation of the time difference using a generalized correlation between a pair of microphone signals [3]. He et al. [4] suggested a deep neural network-based source localization algorithm in the azimuth direction for multiple sources. This approach focused on estimating an incoming direction of a sound and did not localize the actual position of the source.

Recently, many techniques have been proposed to estimate the location of a sound source [5], [6], [7]. Sasaki et al. [5] and Su et al. [6] presented 3D sound source localization algorithms using a disk-shaped sound detector and a linear microphone array such as Kinect and PS3 Eye. Misra et al. [7] suggested a robust localization method in noisy environments using a drone. This approach requires the accumulation of steady acoustic signals at different positions, and thus cannot be applied to a transient sound event or to stationary sound detectors.

An et al. [8] presented a reflection-aware sound source localization algorithm that used direct and reflected acoustic rays to estimate a 3D source position in indoor environments. Our approach is based on this work and takes into account diffraction effects to considerably improve the accuracy.

Interactive sound propagation. There is considerable work in acoustics and physically-based modeling to develop fast and accurate sound simulators that can generate realistic sounds for computer-aided design and virtual environments. Geometry acoustic (GA) techniques have been widely utilized to simulate sound propagations efficiently using ray tracing techniques. Because ray tracing algorithms are based on the sound propagation model at high frequencies, low-frequency wave effects like diffraction are modeled separately.

In addition, an estimation of the acoustic impulse response between the source and the listener was performed using Monte Carlo path tracing [10] or a hybrid combination of geometric and numeric methods techniques [11].

Exact methods to model diffraction are based on directly solving the acoustic wave equation using numeric methods like boundary or finite element methods [12], [13] or the BTM model [14] and its extension to higher order diffraction models [15]. Commonly used techniques to model diffraction with geometric acoustic methods are based on two models: the Uniform Theory of Diffraction (UTD) [16] and the Biot-Tolstoy-Medwin (BTM) model [14]. The BTM model is an accurate diffraction formulation that computes an integral

We have evaluated our method in an indoor environment with three different scenarios, which include a stationary and a dynamically moving sound source along an obstacle that blocks the direct line-of-sight from the listener. In these cases, the diffracted acoustic waves are used to localize the position. We combine our diffraction method with reflection-aware SSL algorithm [8] and observe improvements from 1.22m to 0.7m on average and from 1.45m to 0.79m for the NLOS source. Our algorithm can localize a source generating the clapping sound within 1.38m as the worse error bound in a room of dimension $7m \times 7m$ and 3m height.

II. RELATED WORK

In this section, we give a brief overview of prior work on sound source localization and sound propagation.
of the diffracted sound along the finite edges in the time domain [15], [13], [17]. In practice, the BTM model is more accurate, but is limited to non-interactive applications. The UTD model approximates an infinite wedge as a secondary source of diffracted sounds, which can be reflected and diffracted again before reaching the listener. UTD based approaches have been effective for many real-time sound generation applications, especially in complex environments with occluding objects [18], [10]. Our approach is motivated by these real-time simulation and proposes a real-time source localization algorithm using UTD.

III. DIFFRACTION-AWARE SSL

We present our diffraction-aware SSL based on acoustic ray tracing, starting with giving its overview.

A. Overview

Precomputation. Given an indoor scene, we reconstruct a 3D model as part of the precomputation. We use a Kinect and a laser scanner to capture a 3D point cloud representation of the indoor scene. As shown in Fig. 2, the point cloud capturing the indoor geometry information is generated by the SLAM module from raw depth data and an RGB-D stream collected by the laser scanner and Kinect. Next, we reconstruct a 3D mesh map via the generated point cloud. We also extract wedges from the mesh that have angle, between two neighboring triangles, smaller than the threshold, \( \Theta \). The reconstructed 3D mesh map and the wedges on it are used for our diffraction method at runtime.

Runtime Algorithm. We provide an overview of our runtime algorithm as it performs acoustic ray tracing and sound source localization in Fig. 3. Inputs to our runtime algorithm are the audio stream collected by the microphone array, the mesh map reconstructed in the precomputation, and the robot position localized by the SLAM algorithm. Our goal is find the 3D position of the sound source in the environment. Based on those inputs, we perform acoustic ray tracing supporting direct, reflection, and diffraction effects by generating various acoustic rays (III-B). The source position is computed by estimating the convergence region of the acoustic rays (III-C). Our novel component, acoustic ray tracing with diffraction rays, is highlighted by the blue font in Fig. 3.

B. Acoustic Ray Tracing

In this section, we explain how our acoustic ray tracing technique generates direct, reflection, and diffraction rays. At runtime, we first collect the directions of the incoming sound signals from the TDOA algorithm [19]. For each incoming direction, we generate a primary acoustic ray in the backward direction; as a result, we perform acoustic ray tracing in a backward manner. At this stage, we cannot determine whether the incoming signal is generated by one of the states: direct propagation, reflection, or diffraction. We can determine the actual states of these primary acoustic rays while performing backward acoustic ray tracing. Nonetheless, we denote this primary ray as the direct acoustic ray since the primary ray is a direct ray from the listener’s perspective.

We represent a primary acoustic ray as \( r^0_n \) for the \( n \)-th incoming sound direction. Its superscript denotes the order of the acoustic path, where the 0-th order denotes the direct path from the listener. We also generate a (backward) reflection ray once an acoustic ray intersects with the scene information under the assumption that the intersected material mainly consists of specular materials [8]. The main difference from the prior method [8] is that we use a mesh-based representation, while the prior method used a voxel-based octree representation for intersection tests. This mesh is computed during precomputation and we use the triangle normals to perform the reflections. As a result, for the \( n \)-th incoming sound direction, we recursively generate reflection rays with increasing orders, encoded by a ray path that is defined by \( R_n = [r^0_n, r^1_n, ...] \). The order of rays increases as we perform more reflection and diffraction.

C. Handling Diffraction with Ray Tracing

We now explain our algorithm to model the diffraction effects efficiently within acoustic ray tracing to localize the sound source. Since our goal is to achieve fast performance in localizing the sound source, we use the formulation based on Uniform Theory of Diffraction (UTD) [16]. The incoming sounds collected by the microphone array consist of contributions from different effects in the environment, including reflections and diffractions.

Edge diffraction occurs when an acoustic wave hits the edge of a wedge. In the context of acoustic ray tracing, when an acoustic ray hits an edge of a wedge between two neighboring triangles, the diffracted signal propagates into all possible directions from that edge. The UTD model assumes that the point on the edge causing the diffraction effect is an imaginary source generating the spherical wave [16].

In order to solve the problem of localizing the sound source, we simulate the process of backward ray tracing.
Suppose that an $n$-th incoming sound direction denoted by the ray $r_{n}^{j-1}$ is generated by the diffraction effect at an edge. In an ideal case, the incoming ray will hit the edge of a wedge and generate the diffraction acoustic ray $r_{n}^{j}$, as shown in Fig. 4a in the figure of (a), $r_{n}^{(j)}$ is shown. It is important to note that there can be an infinite number of incident rays generating diffractions at the edge. Unfortunately, it is not easy to link the incident direction exactly to the edge generating the diffraction. Therefore, we generate a set of $N_d$ different diffraction rays in a backward manner that covers the possible incident directions to the edge based on the UTD model. This set is generated based on an assumption that one of those generated rays might have the actual incident direction causing the diffraction. When there are sufficient acoustic rays, including the primary, reflection, and diffraction rays, it is highly likely that those rays will pass through or near to the sound source location; we choose a proper value of $N_d$ by analyzing diffraction rays (Sec. IV).

This explanation begins with the ideal case, where the acoustic ray $r_{n}^{j-1}$ hits the edge of a wedge. Because our algorithm works on the real environment containing various types of errors from sensor noises and resolution errors from the TDOA method, it is rare that an acoustic ray intersects an edge exactly.

In order to support various cases that arise in real environments, we propose using the notion of diffraction-condition between a ray and a wedge. The diffraction-condition simply measures how close the ray $r_{n}^{j-1}$ passes to an edge of the wedge. Specifically, we define the diffractability $\nu_d$ according to the angle $\theta_d$ between the acoustic ray and its ideally generated ray for the diffraction with the wedge: i.e. $\nu_d = \cos(\theta_d)$, where the cos function is used to normalize the angle $\theta_d$ (Fig. 5).

Given an acoustic ray $r_{n}^{j-1}$, we define its ideally generated ray $r_{n}^{j-1}$ as the projected ray of $r_{n}^{j-1}$ on the edge of the wedge where the end point $m_d$ of $r_{n}^{j-1}$ is on that edge (refer to the geometric illustration on Fig. 5). The point $m_d$ is located at the position closest to the point $m_{D}^{j-1}$ of the input ray $r_{n}^{j-1}$; due to the page limit, we do not show its detailed derivation, but it can be defined based on our high-level description.

If the diffractability $\nu_d$ is larger than a threshold value, e.g., 0.95 in our tests, our algorithm determines that the acoustic ray is generated from the diffraction at the wedge, and we thus generate the secondary, diffraction ray at the wedge in the backward manner.

We now present how to generate the diffraction rays when the acoustic ray satisfies the diffraction-condition. The diffraction rays are generated along the surface of the cone (Fig. 4a) because the UTD model is based on the principle of Fermat [9]: the ray follows the shortest path from the source to the listener. The surface of the cone for the UTD model contains every set of the shortest paths. When the acoustic ray $r_{n}^{j-1}$ satisfies the diffraction-condition, we compute outgoing directions for those diffraction rays. Those directions are the unit vectors generated on that cone and can be computed on a local domain as shown in Fig. 4b.

\[
\hat{d}_{n}^{(j,p)} = \begin{bmatrix}
\cos(\theta_{/2} + p \cdot \theta_{/2}) \\
\sin(\theta_{/2} + p \cdot \theta_{/2}) \\
-\cos \theta_d
\end{bmatrix}
\sin \theta_{/2} \sin \theta_{/2} \sin \theta_{/2} = 0.95
\]

where $\hat{d}_{n}^{(j,p)}$ denotes the outgoing unit vector of a $p$-th diffraction ray among $N_d$ different diffraction rays, $\theta_{/2}$ is the angle between two triangles of the wedge, $\theta_d$ is the angle of the cone that is same as the angle between the outgoing diffraction rays and the edge on the wedge, and $\theta_{/2}$ is the offset angle between two sequential diffraction rays, i.e. $\hat{d}_{n}^{(j,p)}$ and $\hat{d}_{n}^{(j,p+1)}$, on the bottom circle of the cone.

Given a hit point $m_d$ by an acoustic ray $r_{n}^{j-1}$ on the wedge, we transform the outgoing directions in the local space to the world space by aligning their coordinates ($\hat{e}_x, \hat{e}_y, \hat{e}_z$). Based on those transformed outgoing directions, we then compute the outgoing diffraction rays, $\hat{r}_{n}^{(j)} = \{r_{n}^{(j)}, ..., r_{n}^{(j,N_d)}\}$, starting from the hit point $m_d$.

In order to accelerate the process, we only generate the diffraction rays in the shadow region, which is defined by the wedge; the rest of the shadow region is called the illuminated region. We use this process because covering only the
shadow region but not the illuminated region generates minor errors for a simulation of the sound propagation [18].

Given the new diffraction rays, we apply our algorithm recursively and generate another order of reflection and diffraction rays. Given the \( n \)-th incoming direction signal, we generate acoustic rays, including direct, reflection, and diffraction rays and maintain the ray paths \( R_n \) in a tree data structure. The root of this tree represents the direct acoustic ray, starting from the microphones. The depth of the tree denotes the order of its associated ray. Note that we generate one child and \( N_d \) children for handling reflection and diffraction effects, respectively.

D. Estimating the Source Position

We explain our method used to localize the sound source position using acoustic rays. Our estimation is based on Monte-Carlo localization (MCL), also known as the particle filter [8]. Our estimation process assumes that there is a single sound source in the environment, which causes a high probability that all those acoustic ray paths pass near that source; Handling multiple targets using a particle filter has been also studied [20]. In other words, the acoustic rays converge in a region located close to the source, and our estimation aims to identify such a convergence region out of all the generated rays.

The MCL approach generates initial particles in the space as an approximation to the source locations. It allocates higher weights to particles that are closer to acoustic rays and re-samples the particles to get more particles in regions with higher weights [8]. Specifically, we adopt the generalized variance, which is a one-dimensional measure for multi-dimensional scatter data, to see whether particles have converged. When the generalized variance is less than a threshold (e.g., \( \sigma_e = 0.5 \)), we treat that a sound occurs and the mean position of those particles as the estimated sound source position.

IV. RESULTS AND DISCUSSION

In this section, we describe our setup consisting of a robot with microphones and testing environments, and highlight the performance of our approach. The hardware platform is based on Turtlebot2 with a 2D laser scanner, Kinect, a computer with an Intel i7 process, and a microphone array, which is an embedded system for streaming multi-channel audios [21], consisting of eight microphones. For all the computations, we use a single core, and perform our estimation every 200ms, supporting five different estimations in one second.

Benchmarks. We have evaluated our method in indoor environments containing a box-shaped object that blocks direct paths from the sound to the listener. We use two scenarios: a stationary sound source and a moving source. As shown in Fig. 6 we place an obstacle between the robot and the stationary sound source, such that the source is not in the direct line-of-sight of the robot (i.e. NLOS source). We use another testing environment with a source moving along the red trajectory, as shown in Fig. 1a. These two scenarios are tested on the same room that size is \( 7m \times 7m \) and 3m height.

During the precomputation phase, we perform SLAM and reconstruct a mesh of the testing environment. We ensure that the resulting mesh has no holes using the MeshLab package.

Stationary sound source with an obstacle. We evaluate the accuracy by computing the L2 distance errors between the positions estimated by our method and the ground-truth positions. We use two types of sound signals: the clapping sound and male speech, where male speech has more low-frequency components than the clapping sound (dominant frequency range of the clapping sound: 2k–2.5kHz, and of male speech: 0.1k–0.7kHz).

We compare the accuracy of our approach with that of Reflection-Aware SSL (RA-SSL) [8], which models direct sound and indirect reflections, but no diffraction. For the stationary source producing clapping sound (Fig. 7a), the average distance errors of the RA-SSL and our method are 1.4m and 0.6m, respectively. There are configurations of the sound source that are not visible to the microphone (NLOS). In this case, we observe 130% accuracy by modeling these diffraction rays.

Fig. 7b shows the localization accuracy for the male speech signal, which has more low-frequency components. The measured distance errors are, on average, 1.12m for RA-
SSL and 0.82m for our approach. While we also observe meaningful improvement, it is less than we see with the clapping sound. Our method supports diffraction, but diffuse reflection is not yet supported. Given the many low-frequency components of male speech, we observe that it is important to support diffuse reflection in addition to diffraction. Nonetheless, by modeling diffraction for the male speech, we observe meaningful improvement (37% on average) in localization accuracy.

Moving sound source around an obstacle. We also evaluated our algorithm on a more challenging environment that contains a sound source (clapping sound) moving along the red trajectory shown in Fig. 1a. Its accuracy graphs are presented in Fig. 1b. Its accuracy graphs are presented in Fig. 1b. The average distance errors of RA-SSL and our method are 1.15m and 0.95m, respectively, in the case of the static source with clapping sound. Also, the average running times for acoustic ray tracing and particle filter are 0.09ms and 72ms; our unoptimized particle filter uses 100 particles and computes weights of them against all the other rays. When we are done on estimating the location within the time budget, we let our process to be in the idle state.

V. CONCLUSIONS & FUTURE WORK

We have presented a novel diffraction-aware source localization algorithm. Our approach can be used for localizing a NLOS source and models the diffraction effects using the uniform theory of diffraction. We have combined our method with indirect reflections and have tested our method in various scenarios with static and moving sound sources with different sound signals.

While we have demonstrated benefits of our approach, it some limitations. The UTD model is an approximate model and mainly designed for infinite wedges. As a result, its accuracy may vary in different environment. We observed lower accuracy for low-frequency sounds (male voice), mainly due to the diffuse effect. Our implemented approach is limited to a single sound source in the environment and does not model all the scattering effects. As part of future work, we would like to address these problems.

Analysis of diffraction rays. By modeling the diffraction effects, we increase the number of generated rays, resulting in a computational overhead. As a result, we measure the average accuracy error and computation time as a function of $N_d$, the number of diffraction ray for simulating each edge diffraction. As shown in Fig. 8, the average accuracy error gradually decreases, but we found that when $N_d$ is in a range of 2 to 5, the accuracy is rather saturated. Since we can accommodate to use up to $N_d = 5$ given our runtime computation budget (0.2 s), we use $N_d = 5$ across all the experiments. In this case, the average numbers of direct, reflection, and diffraction rays are 18, 26, and 184, respectively, in the case of the static source with clapping sound. Also, the average running times for acoustic ray tracing and particle filter are 0.09ms and 72ms; our unoptimized particle filter uses 100 particles and computes weights of them against all the other rays. When we are done on estimating the location within the time budget, we let our process to be in the idle state.

Fig. 7. These graphs compare the localization distance errors of our method with the prior, reflection-aware SSL method [8] with the clapping sound source (a) and male speech signal source (b); green regions indicate no sound in that period. The average distance errors of RA-SSL and our method are 1.4m and 0.6m in (a), and 1.12m and 0.82m in (b), respectively. The use of diffraction considerably reduces the localization errors.

Table I

| Scenario        | Stationary* | Stationary* | Dynamic* | Dynamic* |
|-----------------|-------------|-------------|----------|----------|
| Sound           | Clapping    | Male voice  | Clapping | Clapping |
| RA-SSL          | 1.4m        | 1.12m       | 1.15m    | 1.83m    |
| Ours            | 0.6m (130%) | 0.82m(37%)  | 0.7m(64%)| 0.95m(92%)|

Fig. 8. This figure shows the average accuracy error and computation time for our method on an Intel i7 processor 6700, as a function of $N_d$ the number of diffraction rays generated for simulating the edge diffraction.
REFERENCES

[1] Craig C Douglas and Robert A Lodder, “Human identification and localization by robots in collaborative environments”, Procedia Computer Science, vol. 108, pp. 1602–1611, 2017.

[2] Muhammad Imran, Akhtar Hussain, Nasir M Qazi, and Muhammad Sadiq, “A methodology for sound source localization and tracking: Development of 3d microphone array for near-field and far-field applications”, in Applied Sciences and Technology (IBCAST), 2016 13th International Bhurban Conference on. IEEE, 2016, pp. 586–591.

[3] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay”, IEEE Trans. Acoust., Speech, Signal Process., vol. 24, no. 4, pp. 320–327.

[4] Petr Motlicek Weipeng He and Jean-Marc Odobez, “Deep neural networks for multiple speaker detection and localization”, in ICRA, 2018.

[5] Y. Sasaki, R. Tanabe, and H. Takemura, “Probabilistic 3d sound source mapping using moving microphone array”, in IROS, 2016.

[6] D. Su, T. Vidal-Calleja, and J. V. Miro, “Towards real-time 3d sound sources mapping with linear microphone arrays”, in ICRA, 2017.

[7] Pragyan Mohapatra Prasant Misra, A. Anil Kumar and Balamuralidhar P., “Droneears: Robust acoustic sound localization with aerial drones”, in ICRA, 2018.

[8] Inkyu An, Myungbae Son, Dinesh Manocha, and Sung-eui Yoon, “Reflection-aware sound source localization”, in ICRA, 2018.

[9] Joseph B Keller, “Geometrical theory of diffraction”, JOSA, vol. 52, no. 2, pp. 116–130, 1962.

[10] Carl Schissler, Ravish Mehra, and Dinesh Manocha, “High-order diffraction and diffuse reflections for interactive sound propagation in large environments”, ACM Transactions on Graphics (TOG), vol. 33, no. 4, pp. 39, 2014.

[11] Hengchin Yeh, Ravish Mehra, Zhimin Ren, Lakulish Antani, Dinesh Manocha, and Ming Lin, “Wave-ray coupling for interactive sound propagation in large complex scenes”, ACM Transactions on Graphics (TOG), vol. 32, no. 6, pp. 165, 2013.

[12] B Teng and R Eatock Taylor, “New high-order boundary element methods for wave diffraction/radiation”, Applied Ocean Research, vol. 17, no. 2, pp. 71–77, 1995.

[13] Sara R Martin, U Peter Svensson, Jan Slechta, and Julius O Smith, “A hybrid method combining the edge source integral equation and the boundary element method for scattering problems”, in Proceedings of Meetings on Acoustics 171ASA. ASA, 2016, vol. 26, p. 015001.

[14] U Peter Svensson, Roger I Fred, and John Vanderkooy, “An analytic secondary source model of edge diffraction impulse responses”, The Journal of the Acoustical Society of America, vol. 106, no. 5, pp. 2331–2344, 1999.

[15] Andreas Asheim and U Peter Svensson, “An integral equation formulation for the diffraction from convex plates and polyhedra”, The Journal of the Acoustical Society of America, vol. 133, no. 6, pp. 3681–3691, 2013.

[16] Robert G Kouyoumjian and Prabhakar H Pathak, “A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface”, November, vol. 88, pp. 1448–1461, 1974.

[17] Lakulish Antani, Anish Chandak, Micah Taylor, and Dinesh Manocha, “Efficient finite-edge diffraction using conservative from-region visibility”, Applied Acoustics, vol. 73, no. 3, pp. 218–233, 2012.

[18] Nicolas Tsingos, Thomas Funkhouser, Addy Ngan, and Ingrid Carl-bom, “Modeling acoustics in virtual environments using the uniform theory of diffraction”, in Proceedings of the 28th annual conference on Computer graphics and interactive techniques. ACM, 2001, pp. 545–552.

[19] J.-M. Valin, F. Michaud, and J. Rouat, “Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering”, Robot. Auton. Syst., vol. 55, no. 3.

[20] K. Okuma, A. Taleghani, N. d. Freitas, J. J. Little, and D. G. Lowe, “A boosted particle filter: Multitarget detection and tracking”, in ECCV, 2004.

[21] S. Briere, J.-M. Valin, F. Michaud, and D. Létourneau, “Embedded auditory system for small mobile robots”, in ICRA, 2008.