Passive acoustic non-line-of-sight localization without a relay surface

Tal I. Sommer[1,2] and Ori Katz[1,*]

[1] *Department of Applied Physics, Hebrew University of Jerusalem, Jerusalem 9190401, Israel*

[2] *Alexander Grass Center for Bioengineering, Hebrew University of Jerusalem, Jerusalem 9190401, Israel*

(Received 5 June 2025; revised 26 November 2025; accepted 14 January 2026; published 20 February 2026)

The detection and localization of a source hidden outside the line-of-sight (LOS) traditionally rely on the acquisition of indirect signals, such as those reflected from visible relay surfaces such as floors or walls. These reflected signals are then utilized to reconstruct the obscured scene. In this study, we present an approach that utilizes signals diffracted from an obstacle edge to achieve three-dimensional localization of an acoustic point source situated outside the LOS. The key to the present method is that it removes the dependency on a relay surface for localization. We address two scenarios—a doorway and a convex corner—and propose a localization method for each of them. For the first scenario, we utilize the two edges of the door as virtual detector arrays. For the second scenario, we exploit the spectral signature of a knife-edge diffraction, inspired by the human perception of sound location by the head-related transfer function. In both methods, knife-edge diffraction is utilized to extend the capabilities of non-line-of-sight (NLOS) acoustic sensing, enabling localization in environments where conventional relay-surface based approaches may be limited.

DOI: 10.1103/p97k-sf71

I. INTRODUCTION

The challenge of non-line-of-sight (NLOS) imaging has interesting applications in diverse fields, including autonomous navigation, remote surveillance, search and rescue, and more. Obtaining a direct LOS in these scenarios may be unsafe, challenging, or even impossible. Addressing the challenge involves reconstructing the location and potentially also the shape of objects hidden from direct view, such as those around corners or behind barriers [1].

In the past decade, there has been a growing interest and research focus on optical and acoustical techniques, both active and passive, that perform NLOS imaging, with impressive results in a variety of scenarios. Both optical and acoustical-based techniques seek to address this challenge by extracting information from indirect reflected signals that reach the detection system after being either specularly or diffusively reflected from a visible relay surface, such as a wall, a floor, or a ceiling. While active NLOS imaging systems, both optical [1–4] and acoustical [5] exploit a relay surface to reflect controlled emitted signals to the target object and back, passive systems [6–12] detect reflected signals that are emitted from the target source itself to infer location. In this paper, we will discuss passive localization systems.

Traditionally, NLOS localization relies on measuring the time of arrival (TOA) and/or intensity distribution *on a visible relay surface* [4,7,8,13] to determine the position and shape of targets hidden from view. In the case of TOA measurements, every point on the relay surface effectively serves as a virtual detector and the target can be localized in three dimensions (3D) from this two-dimensional (2D) "array" of virtual detectors [1,14]. In the case of intensity-only measurements, the shadow of the edge of the barrier is utilized to infer the source location. When the barrier has only one edge, as in a convex corner, localization information is available only in one axis [7].

These methods all rely on the presence of a relay surface. If a relay surface is not present, or if the reflections from the relay surface do not reach the detection system, e.g., as in the case of specular mirrorlike reflections in acoustics, one is limited to utilizing the signals that arrive directly from the visible edges of the barrier or corner. Such "direct" signals indeed exist, as the sharp edge of the barrier diffracts the signal from the source in what is known as knife-edge diffraction [15–17].

The method proposed here hinges on the *removal of the dependency on a visible relay surface* for localization.

*Contact author: orik@mail.huji.ac.il

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

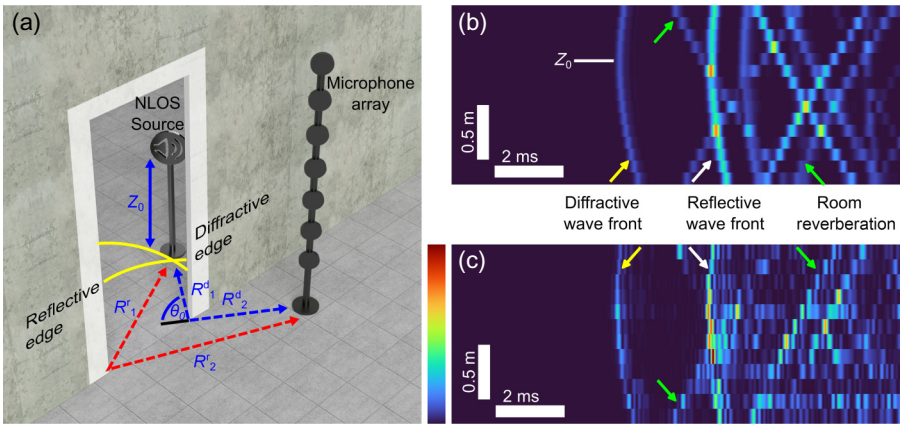


FIG. 1. (a) A conceptual illustration of NLOS localization in a doorway scenario. (b),(c) The amplitude envelope of the signal detected by the microphone array in this scenario, $A(z,t)$: (b) simulation; (c) experiment. The detected signals from the hidden acoustic point source are the result of two phenomena (depicted by arrows): the spherical wave front that interacts with the door edge that is closer to the microphone array undergoes diffraction, which redirects the wave toward the microphone array. This diffracted wave front (yellow arrows) is the first to be detected by the array. The spherical wave front that interacts with the farther edge of the door is reflected toward the microphone array. This reflected wave front (white arrows) is the second to be detected by the array. Additionally, later-arriving reverberations echo throughout the room (green arrows). The color-bar range in (b) and (c) is [0–1].

Accordingly, we investigate two common scenarios of passive acoustic NLOS localization that operate without a relay surface: a doorway and a single edge. In the doorway scenario [Fig. 1(a)], the two edges of a doorway act effectively as two virtual line arrays, where the signal from one edge is relayed through diffraction and from the other edge through both diffraction and reflection. These two virtual arrays can be utilized as the conventional 2D relay surface used in previous works. In contrast, as we discuss below, in the case of a single edge that acts as a single virtual line array (Fig. 3), additional information is required to assess the azimuth of the source.

In this work, we address passive acoustic NLOS source localization in these two scenarios. Using an array of microphones, we measure the acoustic field on the detection side that originates from a hidden point-source emission. In the case of a doorway [Fig. 1(a)], we calculate the TOA of the measured signals at two visible edges of the door. The TOA from the two line arrays is then used as in a traditional NLOS TOA reconstruction [1,2] to localize the source in 3D. In a single-edge scenario, as is encountered in a single barrier edge (Fig. 3) or a convex corner, we similarly utilize the calculated TOAs at the single edge, which allow us to localize the range and height of the source. The azimuthal angle, θ, is determined from the relative spectral signature of the knife-edge diffraction [15] measured by two microphone arrays, exploiting the fact that lower frequencies diffract more efficiently at large diffraction angles. The use of a spectral signature to determine location is similar to the human perception of 3D acoustic location via the head-related transfer function (HRTF)[18–20].

II. SCENARIO 1—A DOORWAY

We first consider the doorway scenario, as illustrated in Fig. 1(a). In this scenario, a hidden source emits a short acoustic pulse. Part of the acoustic waves from the source is diffracted in a knife-edge diffraction from the closer edge of the door to a microphone array [dashed blue curve in Fig. 1(a)] and another part is reflected from the farther edge of the door to the microphone array (dashed red curve). The challenge is to utilize these two received waves to localize the hidden source in 3D.

We begin our study with a numerical analysis of this scenario using a finite-difference time-domain (FDTD) simulator for propagating acoustic waves (the *k*-Wave toolbox for the MATLAB software package [21,22]; see Sec. A of the Supplemental Material [23]) to simulate the propagation of acoustic fields in the scene, and the resulting signals expected to be measured by the microphone array. The simulated 3D scene geometry is depicted in Fig. 1(a) [and in a top view in Fig. 2(b)]. It consists of a vertical array of 15 microphones placed on a rod at a distance of $R^d_2 = 0.8$ m from the diffractive edge of a doorway. The walls, ceiling, and floor were simulated as high speed of sound and their density compared to air (for full simulation parameters, see Sec. A of the Supplemental Material [23]).

One example result of the simulated waves at the microphone array for an object located at $R^d_1 = 3.2$ m, $\theta_0 = 25^\circ$, $z_0 = 1.5$ m is presented in Fig. 1(b). The plotted values are the envelope amplitude after bandpass filtering at between 500 Hz and 9 kHz and amplitude demodulation, at each microphone height, z, over time, t: $A(z,t)$. As expected, the received signals are comprised of two

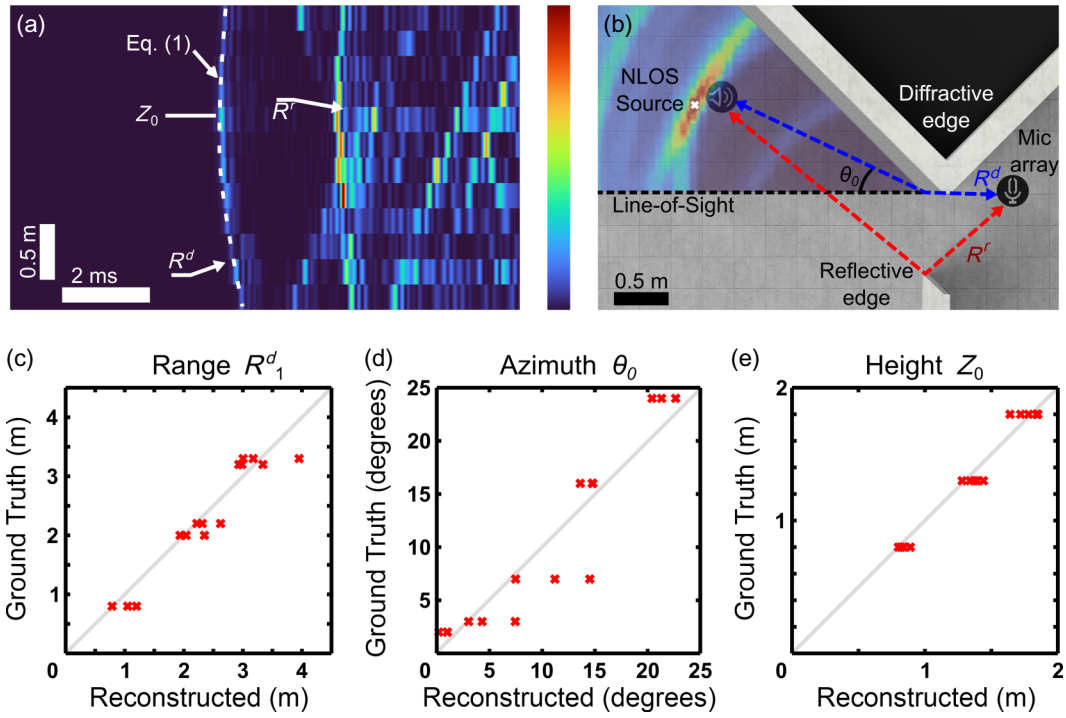


FIG. 2. (a) The amplitude envelope of the signal detected by the microphone array in a doorway scenario, $A(z,t)$. The spherical wave-front attributes of Eq. (1) are depicted on the measured signal. (b) A top view of the scene and the localization heat map reconstructed from the data in (a). (c) The reconstructed range between the source and the microphone array shows a root-mean-squared error (RMSE) of approximately 18%. (d) The reconstructed azimuth angle between the LOS and the microphone array shows an RMSE of approximately 3°. (e) The reconstructed height of the source shows a mean prediction error of approximately 6%. The color-bar range in (a) and (b) is [0–1].

types of waves and reverberations. The three types of waves are depicted by arrows in Figs. 1(b) and 1(c): the diffracted wave from the closer edge of the doorway (yellow arrow), the reflected wave from the farther edge of the doorway (white arrow), and reverberating waves following each of the above (green arrows) that originate from multiple reflections from other surfaces or objects in the scene.

Considering a point source, we expect the detected wave front to be spherical, where the radius of curvature is defined by the propagation distance from the source, i.e., $R_1^d + R_2^d$, and $R_1^r + R_2^r$. The diffracted wave front has a radius of curvature that is shorter than the reflected wave-front radius of curvature [see Figs. 2(a) and 2(b)]. Explicitly, $R_1^d + R_2^d \equiv R^d < R^r \equiv R_1^r + R_2^r$, where R_2^d and R_2^r are known scene constants [propagation distances from the edges to the detection array; see Fig. 1(a)] and R_1^d and R_1^r are the unknown parameters of the hidden source location.

Ignoring the reverberations in the background, in each microphone, each of the two wave fronts (diffracted of reflected) is manifested as a single TOA. For a hidden point source located at a height z_0 above the floor and at a propagation distance (R^r or R^d) from the microphone array, emitting a short pulse at an unknown time t_0, the wave-front time of arrival, TOA(z), is given by

$$\mathrm{TOA}(z) = t_0 + \sqrt{(z - z_0)^2 + R_0^2}, \tag{1}$$

where: (i) R_0 will be equal to either R^d or R^r for either the diffracted or reflected wave, and will be manifested in the curvature of the detected wave front; (ii) the height of the source above the ground, z_0, will be also the height at which the detected signal will first arrive [see the z_0 mark in Fig. 1(b)]; and (iii) the unknown emission time, t_0, will cause a temporal shift of the detected waves. Note that R_0 also causes a temporal shift of the detected waves, but unlike t_0, will also change the detected wave-front curvature.

We used Eq. (1) to localize the source in 3D in three steps. First, we found the TOA of the diffracted wave front in each microphone. This was done by finding the first peak of the signal that was significantly stronger than the mean background amplitude.

Second, we fitted the found TOA of the diffracted wave front to Eq. (1) [Fig. 2(a)], yielding R^d, t_0, and z_0.

Knowing z_0 and t_0 from the initial fit, we can now use a modified beam-forming algorithm using the two edges as a 2D relay surface, to reconstruct the initial amplitude of the source at each point in the hidden scene, as we explain

below. An example of the beam-forming-reconstructed amplitude metric from one experimental measurement is presented as a heat map in Fig. 2(b). The location of the source is found from the maximum-amplitude metric location.

The reconstruction heat map is generated according to the following.

The metric, M, for each point is calculated using the summation of the absolute value of the signal envelope under the restrictions of the mask:

(I) For each spatial position in the reconstructed grid (x, y, z), the horizontal propagation distances to the microphone array $R^{d,r} = \sqrt{(x - x_{d,r})^2 + (y - y_{d,r})^2} + R_2^{d,r}$ were calculated.

(II) For each of the two distances R^d and R^r, Eq. (1) was used to provide two TOAs for all microphones: $TOA_r(z)$ and $TOA_d(z)$.

(III) For each of the microphones, the amplitude in the trace $A(z, t)$ at a temporal gate of length $0.2ms$ around each of the two TOAs was averaged, providing two scalar numbers that provide an estimate for the detected amplitudes of the reflected and diffracted waves: A_r and A_d, respectively.

(IV) A metric, M, for the current point was calculated by

$$M\left(R^d, R^r\right) = \frac{A_d}{(R^d)^2}\frac{A_r}{(R^r)^2} \tag{2}$$

where the multiplicative weighting by $\left(1/R^d R^r\right)^2$ was chosen to prioritize closer locations and to reduce the effect of strong late-arriving reverberations. In Fig. 2(b), we display the reconstructed metric from one experimental measurement performed as explained below.

The above reconstruction is in fact a delay-and-sum beam forming of the detected amplitudes, ignoring the detected phases, i.e., the equivalent of using the two edges of the door as the relay surface. We found empirically that this gives better results compared to conventional delay and sum on the measured data sets.

Following the numerical study, we performed a set of experiments in a large concrete-walled basement with the geometry given in Fig. 2(b). As the hidden source, we used the hand claps of a test subject at various positions hidden outside the LOS behind the doorway [the colored area in Fig. 2(b)]. The spectrum of the hand claps was significant above the background noise level mostly in the frequency range from 500 Hz to 9 kHz (see Fig. S1 in the Supplemental Material [23]). The signals were recorded by a vertical array of 15 microphones with 13-cm spacing between neighboring microphones, starting at a height of 46 cm above the floor, up to 228 cm above the floor. The distance of the array from the diffractive edge was $0.8m$ and the width of the doorway was 0.9 m. Fifteen different source locations were tested. For each test, a single hand clap was measured and the position of the source was reconstructed as explained above. The results for the reconstructed positions versus the real source positions are presented in Figs. 2(c)–2(e). Each of the panels presents one coordinate of the source position: its distance from the diffractive edge (R_1^d), its azimuth relative to the LOS (θ_0), and its height (z_0).

Note that the reconstructed location is sensitive to errors in the estimation or measurement of the geometry of the visible part of the scene. In particular, the measured distance between the array and the edges (R_2^d and R_2^r) causes an additive constant bias to the azimuth. The results presented in Fig. 2(b) were obtained after a small change of 2–8 cm in the measured scene parameters, which is within our experimental measurement error.

Inspecting the results, for our experimental parameters, the reconstructed range root-mean-squared error (RMSE) is approximately 18% of the true range. The reconstructed azimuth RMSE is approximately 3° and the reconstructed height RMSE is approximately 6%, which is at worst less than the 13-cm spacing between the microphones.

III. SCENARIO 2—A SINGLE KNIFE-EDGE

Following the completion of the empirical study of the doorway scenario, we turned to consider the scenario of NLOS localization in a semi-infinite aperture when only a single edge is visible, as illustrated in Fig. 3, as is the case when the source is hidden beyond e.g., a corner of a building. In this case, the signal arriving at the microphone array will only include the diffracted wave front (and the latter reverberations). Thus, while the diffracted wave front measured in a single vertical array can provide the range and height of the source, additional information is required to estimate its azimuth θ_0. In our results below, we estimate the azimuth from the spectral ratio of the diffracted waves as measured by two vertical arrays (Fig. 3). The two arrays allow us to exploit the fact that lower frequencies have stronger diffraction amplitude at large angles, resulting in an azimuth-dependent spectrum (depicted as colored lines in Fig. 3).

Theoretically, the diffracted amplitude from a semi-infinite aperture, also known as knife-edge diffraction, is dependent on the hidden source azimuth, θ, the frequency, f, and the geometry of the scene [15]. These parameters are encompassed in a single parameter, ν, termed the Fresnel-Kirchoff parameter [15]: $\nu \propto \theta\sqrt{f}$.

The diffraction amplitude is commonly characterized by the ratio between the detected diffracted NLOS signal ($A_{knife-edge}$) and the signal that would have been detected at the same position in a free-space configuration ($A_{free-space}$). This ratio is termed the "diffraction loss" (or gain), $\mathscr{L}$ [15–17]. To provide an intuitive insight into the expected result, we review the common simple case of knife-edge

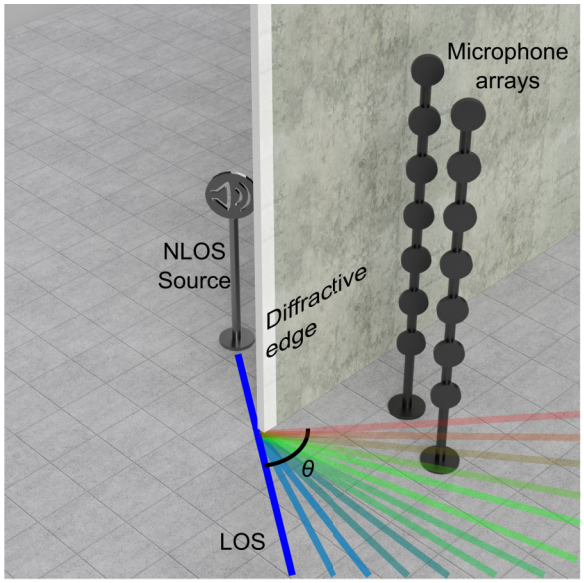


FIG. 3. A conceptual illustration of NLOS localization in a knife-edge scenario. The spherical wave front propagates from the hidden point source and interacts with the edge, resulting in diffraction toward the microphone array. The diffraction loss introduced by this interaction is dependent on both the angle θ between the microphone array and the line of sight (LOS), and on the signal frequency (depicted by different colors). Lower frequencies diffract more efficiently at high azimuths.

diffraction in 2D under the Fresnel approximation (see Sec. B if the Supplemental Material [23]). For this simplistic case, $\mathscr{L}$ can be numerically calculated using the complex Fresnel integral [15]

$$\mathscr{L}(\nu) \equiv \frac{A_{\text{knife-edge}}}{A_{\text{free-space}}} = \frac{1+j}{2}\int_{\nu}^{\infty} \exp\left(-j\frac{\pi}{2}t^2\right)dt. \quad (3)$$

In Figs. 4(a) and 4(b), we present the theoretical diffraction loss according to Eq. (3) of the simple 2D case, for different frequencies and azimuths. As expected, the diffraction loss is larger at larger angles [Fig. 4(a)] and at higher frequencies [Fig. 4(b)].

In Fig. 4(b), we show the dependence of the measured NLOS spectrum on the hidden source azimuth, θ. In the case that the emission spectrum at the source position is known, a single measurement of the diffracted spectrum can yield the source azimuth using Fig. 4(b). Alas, in many realistic scenarios, the source emission spectrum is not known. Thus, the spectral loss is unattainable from a single measurement, as the measured spectrum is the product of the unknown emitted spectrum and the diffraction loss. To overcome this limitation, we measure the diffracted NLOS signal with *two* microphone arrays positioned at different azimuths from the diffractive edge (Fig. 3). We found that, in both theory and experiments, the spectral ratio between the two arrays allows the retrieval of the diffraction loss while cancelling any dependency on the source spectral signature [Fig. 4(c)]. Intuitively, the ratio of the spectral amplitudes measured in two microphone arrays at different azimuths gives, for each frequency, f, the ratio between the diffraction losses at these two azimuths: $\mathscr{L}_{\theta}(f)/\mathscr{L}_{\theta+\Delta\theta}(f)$.

This method is analogous to the human perception of an acoustic source localization via spectral signatures using the HRTF [18–20]. While determining if a source is located to the left or the right of the listener is rather straightforward by measuring the relative TOAs (and/or amplitudes) of the signals received in the left versus the right ear, determining the height of a source or if it is located in front or at the back of the listener cannot be determined from TOAs alone. However, the spectral

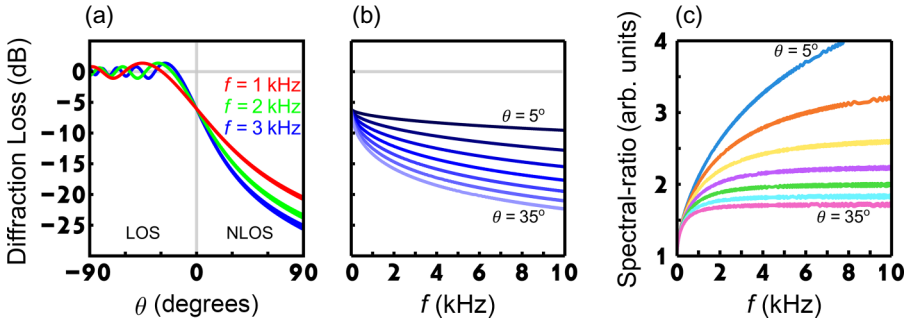


FIG. 4. The theoretical knife-edge diffraction loss for a simplistic 2D scenario (see Sec. B of the Supplemental Material [23]). (a) Knife-edge diffraction introduces a loss that depends on the diffraction angle (θ, as defined in Fig. 3), and the frequency. This panel shows numerical calculations of the induced loss in a specific scenario. (b) In the frequency domain, the knife-edge loss exhibits a unique spectral signature corresponding to the azimuth θ of the source. This panel displays numerical calculations of the spectral signature for a hidden source located at azimuths ranging from 5° to 35°. (c) Direct measurement of the loss spectrum without including the spectrum of the source is not possible. To isolate the loss effect, two microphone arrays positioned at an angular distance of $\Delta\theta = 25°$ between them are used to measure the diffracted signal. The spectral ratio between the two arrays reveals the impact of the diffraction loss on the spectrum, independently. This panel shows numerically calculated spectral ratios for a hidden source at azimuths ranging from 5° to 35°. (a) $\mathscr{L}_f(\theta)$; (b) $\mathscr{L}_{\theta}(f)$; (c) $\mathscr{L}_{\theta}(f)/\mathscr{L}_{\theta+\Delta\theta}(f)$.

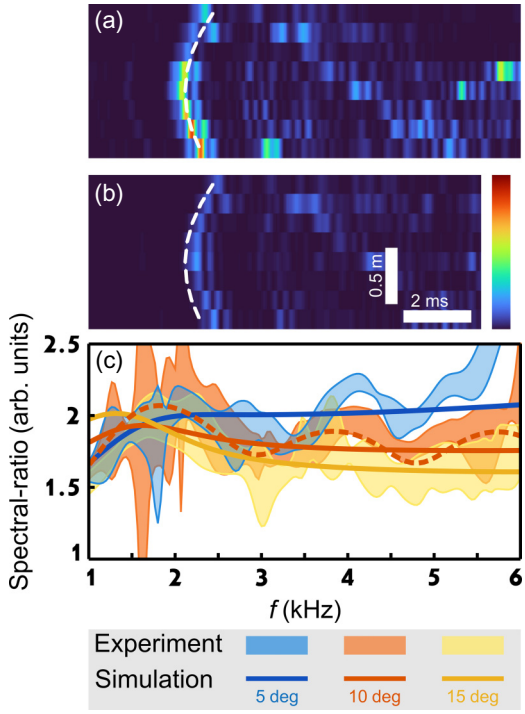


FIG. 5. (a),(b) The experimental-amplitude envelope of the signals recorded by the two microphone arrays in a knife-edge scenario experiment (Fig. 3). The dashed lines indicate the detected first-arriving wave fronts, the arrival times of which are used for retrieving the height and distance of the source. The spectral ratio between the first-arriving wave fronts in the two arrays is used to determine the source azimuth. (c) The spectral ratios of the first-arriving wave fronts between the two microphone arrays. The colored areas display the experimental one standard deviation across all valid measurements. The solid lines show the numerically simulated spectral ratio obtained from an FDTD simulation of acoustic wave propagation in the scene that includes reflections from the walls, floor, and ceiling. The dashed line shows the same as the solid lines but also including the body of the hand-clapping human subject, demonstrating that the source for the spectral modulations consists of reflections from the human subjects. The color-bar range in (a) and (b) is [0–1].

transfer function that arises from the propagation of the signal from a source location to the ear is dependent of the source height due to diffraction from the earlobes, head, and shoulders. This set of HRTFs affects the spectral signature of each source and allows acoustic localization in 3D [18–20].

To validate our approach, we performed a set of experiments in a large concrete-walled basement with a convex corner. As the hidden source, we used the hand claps of a test subject at various positions hidden outside the LOS. The spectrum of the hand claps was significant above the background noise level, mostly in the frequency range from 500 Hz to 9 kHz (Fig. S1 in the Supplemental Material [23]). The signals were recorded by two vertical arrays of eight microphones each (Fig. 3), with a spacing of 26 cm between neighboring microphones, starting at a height of 46 cm above the floor, up to 228 cm above the floor. The distance of both of the arrays from the diffractive edge was 0.8 m and their azimuths were different by 25°. Three source azimuths were tested, all at the same range and height ($R_1^d = 3.1$ m, $z_0 = 1.3$ m). The tested azimuths relative to the LOS, θ_0, were 5°, 10°, and 15°. For each position, ten hand claps were measured. The rather small azimuthal angles were chosen as the spectral ratio metric is more noise sensitive and tolerance sensitive for larger azimuths [above 20°, Fig. 4(c)].

To optimally filter out background noise and room reverberations, we used a short temporal window on the received signals around the first-arriving wave front to calculate the spectral ratio of the first-arriving wave front alone. We found that the spectral ratio is sensitive to the exact choice of window size. In our experiments, the window size was fine tuned until optimal results were produced, at a window size of 0.7 μs.

In Figs. 5(a) and 5(b), we present an example of the amplitude envelope of a single measured signal at the two arrays, $A_{1,2}(z,t)$. The dashed white curves are the fits for the first-arriving wave fronts from the measured TOA. The height and distance of the source were determined from the two curves. The azimuth information was assessed from the spectral ratio, $\mathscr{L}_\theta(f)/\mathscr{L}_{\theta+\Delta\theta}(f)$, calculated directly by the ratio between the power spectra of the signals in the temporal window around the first-arriving wave front. For each measurement, the power spectrum in each pole was calculated from the average of the power spectra of the three central microphones the height of which is closest to the source height.

In Fig. 5(c), we display the experimental spectral ratio for each of the three tested azimuths as colored areas. Each colored area represents the range of one standard deviation of the measured spectral ratios in the experimental measurements of one azimuth. For each azimuth, the spectral ratio was calculated for all valid measurements, where we have disregarded measurements of hand claps that resulted in a relatively low intensity, an undetectable wave front, or a wave-front fit the RMSE of which is larger than 0.1 ms.

In Fig. 5(b), we compare the experimentally measured spectral ratios (colored areas) to numerically simulated spectral ratios (solid-line plots), obtained by FDTD simulation of acoustic wave propagation in the scene, using *k*-Wave [21,22] (the *k*-Wave toolbox for MATLAB; see Sec. A of the Supplemental Material [23]). As expected, at high frequencies, where the spectral ratio is expected to be more significantly different between different azimuth [as seen in Fig. 4(c)], the experimental spectral ratios are indeed close to the numerically simulated ones [Fig. 5(c), solid lines]. The results display additional spectral modulations, which we attribute to interference in the scenes, such as those originating from reflections from the body of the hand-clapping human subject. To confirm this thesis, we added

a simulated spectral ratio of a scene with a human body standing behind the source for one of the scenes [dashed orange line in Fig. 5(c)]. For further discussion and simulations regarding these spectral modulations, see Sec. D of the Supplemental Material [23].

IV. CONCLUSIONS

In conclusion, in this work we have proposed an approach for NLOS localization without a relay surface, utilizing edge diffraction for passive acoustic NLOS localization. This capability allows us to address scenarios that traditional methods, which rely on reflections from visible relay surfaces, cannot address, as we have demonstrated in the single-edged convex corner scenario.

Our numerical results provide a foundation for the proof-of-principle experiments. However, advancements in several areas are required to improve the robustness and applicability of these diffraction-based techniques. In particular, while we have put forward a framework for the analysis of the spectral-ratio-based information in the single-edged convex-corner scenario [Fig. 5(c)], a more advanced signal-processing algorithmic approach, e.g., utilizing neural networks, is likely to improve performance. Our experiments have highlighted the sensitivity of the reconstructed locations to the accuracy of measuring the visible-scene geometry, and the sensitivity to the processed temporal window size. Future work on more advanced algorithms that would reduce these sensitivities is important for deployment in uncontrolled environments.

Important next steps include exploring more complex, dynamic, and noisy environments. Additionally, utilizing nonimpulsive emissions, and emissions from distributed sources, would be interesting extensions to explore.

Overall, our work demonstrates the potential of acoustic diffraction as a valuable source of information for passive NLOS localization, opening interesting avenues for sensing in complex environments.

ACKNOWLEDGMENTS

This project was supported by the H2020 European Research Council (Grant No. 101002406).

T.S. led with regard to data curation, formal analysis, methodology, and the writing and original-draft preparation. O.K. led with regard to supervision and funding acquisition. The authors shared equal responsibility for conceptualization, writing and review, and editing.

DATA AVAILABILITY

The data that support the findings of this paper are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] D. Faccio, A. Velten, and G. Wetzstein, Non-line-of-sight imaging, Nat. Rev. Phys. **2**, 318 (2020).

[2] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging, Nat. Commun. **3**, 745 (2012).

[3] G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio, Detection and tracking of moving objects hidden from view, Nat. Photonics **10**, 23 (2016).

[4] M. O'Toole, D. B. Lindell, and G. Wetzstein, Confocal non-line-of-sight imaging based on the light-cone transform, Nature **555**, 338 (2018).

[5] D. B. Lindell, G. Wetzstein, and V. Koltun, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Long Beach, California, USA, 2019), p. 6780.

[6] R. Czajkowski and J. Murray-Bruce, Two-edge-resolved three-dimensional non-line-of-sight imaging with an ordinary camera, Nat. Commun. **15**, 1162 (2024).

[7] K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Wornell, A. Torralba, and W. T. Freeman, in *Proceedings of the IEEE International Conference on Computer Vision* (Venice, Italy, 2017), p. 2270.

[8] S. W. Seidel, J. Murray-Bruce, Y. Ma, C. Yu, W. T. Freeman, and V. K. Goyal, Two-dimensional non-line-of-sight scene estimation from a single edge occluder, IEEE Trans. Comput. Imaging **7**, 58 (2020).

[9] J. Boger-Lombard, Y. Slobodkin, and O. Katz, Non-line-of-sight passive acoustic localization around corners, Sci. Rep. **13**, 4952 (2023).

[10] J. Boger-Lombard and O. Katz, Passive optical time-of-flight for non line-of-sight localization, Nat. Commun. **10**, 3343 (2019).

[11] O. Katz, P. Heidmann, M. Fink, and S. Gigan, Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations, Nat. Photonics **8**, 784 (2014).

[12] O. Katz, E. Small, and Y. Silberberg, Looking around corners and through thin turbid layers in real time with scattered incoherent light, Nat. Photonics **6**, 549 (2012).

[13] T. Sultan, S. A. Reza, and A. Velten, Towards a more accurate light transport model for non-line-of-sight imaging, Opt. Express **32**, 7731 (2024).

[14] X. Liu, S. Bauer, and A. Velten, Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems, Nat. Commun. **11**, 1645 (2020).

[15] E. C. Jordan, *Electromagnetic Waves and Radiating Systems* (Prentice-Hall, 1958), p. 499.

[16] C. L. Giovaneli, An analysis of simplified solutions for multiple knife-edge diffraction, IEEE Trans. Antennas Propag. **32**, 297 (1984).

[17] G. Millington, R. Hewitt, and F. S. Immirzi, Double knife-edge diffraction in field-strength predictions, Proc. IEE-Part C: Monogr. **109**, 419 (1962).

[18] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, Localization using nonindividualized head-related transfer functions, J. Acoust. Soc. Am. **94**, 111 (1993).

[19] C. I. Cheng and G. H. Wakefield, Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space, J. Audio Eng. Soc. **49**, 231 (2001).

[20] V. Bruschi, L. Grossi, N. A. Dourou, A. Quattrini, A. Vancheri, T. Leidi, and S. Cecchi, A review on head-related transfer function generation for spatial audio, Appl. Sci. **14**, 11242 (2024).

[21] B. E. Treeby and B. T. Cox, *k*-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, J. Biomed. Opt. **15**, 021314 (2010).

[22] B. E. Treeby, J. Budisky, El. S. Wise, J. Jaros, and B. T. Cox, Rapid calculation of acoustic fields from arbitrary continuous-wave sources, J. Acoust. Soc. Am. **143**, 529 (2018).

[23] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/p97k-sf71 for the *k*-Wave simulation parameters, the Fresnel integral simulation parameters, a source and noise spectrum comparison, an investigation regarding the origin of spectral modulations in our experiments, and a MATLAB code for our *k*-Wave simulation.