Obtaining Range Measurements From Ambient Noise Cross-Correlations for the Self-Calibration of Nodes

Joaquín Aparicio[1] and Sverre Holm[2]

[1] *Department of Informatics, University of Oslo, Gaustadalléen 23B, 0377 Oslo, Norway*
[2] *Department of Physics, University of Oslo, Sem Sælands vei 24, 0371 Oslo, Norway*

Abstract

The calibration of nodes in a sensor network or in a positioning system is a time-consuming process. This is especially relevant in indoor environments, where GPS signals are not available. Usually, it involves a user making transmissions either at known points, or by transmitting while moving under the area of interest. Some self-calibration methods have already been proposed for acoustic sensor network systems, that require range calculation between the nodes, to later apply a positioning algorithm. In this work, we evaluated a range mechanism based on the cross-correlation of the ambient noise naturally available indoors, so no active transmissions were required. Results from experimental tests show a minimum Mean Absolute Error (MAE) of 0.93 cm and a maximum MAE of 1.37 cm, when calculating the range between two microphones separated 18.8 cm.

Keywords

Acoustics, indoor positioning, noise, self-calibration, sensor networks

1. Introduction

In certain applications, such as sensor networks or positioning systems, it is necessary to know the position of the nodes that form such systems. For example, sensor network nodes need to know where they are, so the data they measure are linked to a particular location of interest [1]. In positioning systems, usually several nodes act as transmitters (beacons) or receivers. These nodes must have known positions, so the algorithms can calculate the location of a user moving through the environment [2].

Outdoors, the positions of these nodes can be obtained from GPS. However, indoor environments represent a more challenging situation, as GPS signals are highly distorted and attenuated. A calibration process is then conducted to obtain the positions of the beacons. This calibration process can be done either manually, or by obtaining range measurements from a transmitter at known points [3, 4], or at several unknown points spread over the coverage area [5–8]. These methods can be laborious and time-consuming, depending on the number of beacons to locate, and they require actions from the user and the transmission of acoustic signals. Moreover, these measurements have to be repeated in the event that the beacons are moved. Additionally, if the beacons are inadvertently exchanged during a maintenance operation, it could lead to system failure due to the incorrect location of the beacons, or to log incorrect location information in sensor networks, compromising the recorded data.

Therefore, it would be desirable to have an independent, self-calibration process that should be able to obtain the position of the nodes by itself. This approach could also, in principle, detect changes in the infrastructure automatically, and could either annotate and use the new positions, or notify of a potential node swap. One of the earliest self-calibration works was presented in [9]. At least three beacons with known positions broadcasted their location information, together with an ultrasonic signal for range estimation. Other nodes in the network listened to the beacons and could calculate their positions using multi-lateration. Once they knew their position, they could help to locate other nodes in their vicinity. In [10], the self-calibration of a sensor network was obtained by attaching a loudspeaker

to a minimum of five nodes, which acted as transmitters/receivers. After obtaining all pairs of distances between these nodes, their positions were obtained by a Multi-Dimensional Scaling algorithm (MDS), and further optimized by a Levenberg-Marquardt algorithm. Once their positions were known, they could locate other nodes in the network that acted as receivers using a trilateration algorithm. Another example of a self-calibration system was implemented in an outdoor sensor network for source localization in [11]. Each node took turns to transmit a 2048-bit code, while the other array nodes listened and calculated range and bearing information. All measurements were then gathered into a non-linear least squares algorithm that calculated the positions of the nodes. Another approach was presented in [12], where a main node transmitted a unique Complementary Set of Sequences (CSS) initiating the self-calibration process. When this signal was received by the other nodes, they replied with their own unique CSS. These signals were received at the main node, but also at the other ones. The so-called pseudo-Times-of-Flight allowed to calculate the different ranges between nodes, and they were fed to an MDS algorithm, which computed the relative positions of the sensor network. These positions were later refined with a non-linear least squares algorithm.

These systems rely on the active transmission of acoustic signals, which need Line-of-Sight conditions that are not always possible to meet based on nodes placement and directivity, and it uses battery life. A passive alternative is to use the ambient noise already present in the environment. Although it is usually considered an annoyance, noise contains information about the channel in which it propagates, which could be leveraged to obtain the range between nodes. This idea was explored in [13] for the self-calibration of a microphone array for a speech processing application. The distances between pairs of microphones were obtained by assuming a diffuse noise model and applying a model-fitting solution to the coherence equation, which is a function of the distance between the microphones. Further processing by K-means clustering was done to group distance estimates in order to distinguish outliers and select the more robust value. This approach was evaluated for a maximum distance of 20 cm. In order to improve the robustness of this method, in [14] the coherence frames were averaged before the model-fitting step, and a multi-stage clustering was used to remove outliers. Reliable estimations of distances were obtained up to 73 cm. A new model was formulated in [15] assuming diffuse noise conditions, based on the generalized cross-correlation with phase transform (GCC-PHAT). Considering two microphones, the model gave an approximation of the field generated by uncorrelated sources arriving from all possible time-differences of arrival. Distances could be then estimated by a model-fitting solution, minimizing the quadratic error between the model output and the GCC-PHAT output. An experiment was conducted with 8 microphones deployed in a circle with a diameter of 20 cm, and one additional microphone in the center. The reported results improved those of [13], obtaining an error of around 2 cm for the maximum distance of 20 cm.

The approach considered in this work also assumes the noise field to be diffuse, like [13–15]. However, it is not based on a model-fitting solution of the coherence between microphones, but on the processing and direct cross-correlation of the ambient noise recorded at the microphones. This approach was firstly used in geophysics, where it was presented at the end of the 1960s to obtain information from Earth layers [16]. Since then, it has been applied to diverse applications, such as surface-wave tomography [17], volcano monitoring [18], and the self-calibration of an underwater array [19].

We have evaluated here the feasibility of using acoustic ambient noise cross-correlations to obtain ranges in indoor environments. These ranges could be later used in a purely passive beacon self-calibration process for positioning systems and sensor networks. This way the nodes would be able to locate themselves, as well as to detect changes in the infrastructure, without any external aid from a user, and without acoustic emissions, saving time and battery life. The obtained results when calculating the range between two microphones separated 18.8 cm in an office environment had a minimum Mean Absolute Error (MAE) of 0.93 cm and a maximum MAE of 1.37 cm, highlighting the potential of this technique.

The rest of the paper is organized as follows. Section 2 summarizes the fundamentals of this theory. Section 3 describes the experimental setup and signal processing steps, whereas the results are shown in Section 4. Section 5 gathers the conclusions and future work.

2. Fundamentals on the cross-correlation of the noise field

This section presents a summary of the theory behind obtaining the response between two receivers from the cross-correlation of the noise field. The idea originated in geophysics [16], and the demonstration here follows that given in [20], due to its conceptual simplicity. The author proposed the thought experiment in one dimension showed in Fig. 1, composed of two parts: the transmission problem (a), and the earthquake problem (b).
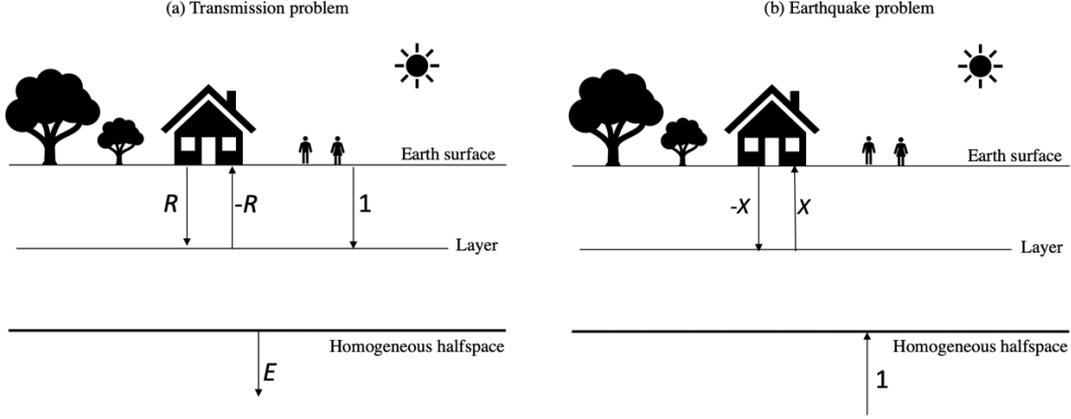
Figure 1: Problem description, showing (a) the transmission problem, and (b) the earthquake problem.

In the transmission problem, an impulse is generated on the surface propagating downwards. Here it is assumed that, for reflection purposes, the Earth surface is a perfect reflector, and the halfspace a perfect transmitter. It is also assumed that the medium is homogeneous. Part of the transmitted signal is reflected at a layer at a certain depth and it bounces back to the surface ($-R$), where it is reflected down again (R). Part of the signal is transmitted through the layer and escapes (E) at the homogeneous halfspace. The net downgoing energy flux can be written then as:

$$\bar{D}D - \bar{U}U,\tag{1}$$

where $\bar{D}D$ is all the energy going down, $\bar{U}U$ is all the energy going up, and the bar indicates complex conjugate. At the Earth surface, $D = 1 + R$ and $U = -R$, whereas at the halfspace, $D = E$ and $U = 0$.

Assuming that there is no absorption, and applying the conservation of energy flux at the Earth surface and the halfspace, we get to (2):

$$\overline{(1 + R)}(1 + R) - \bar{R}R = \bar{E}E\tag{2}$$

After rearranging terms:

$$\bar{R} + 1 + R = \bar{E}E\tag{3}$$

Which means that we should put a receiver in the halfspace to measure the escaping wave in order to obtain information about the layer. The Earthquake problem shown in Fig. 1 (b) represents the principle of reciprocity of the transmission problem: now the emitter is at the halfspace, and we measure the signals at the surface, X. Applying the reciprocity principle, $\bar{E}E = \bar{X}X$ and (3) becomes:

$$\bar{R} + 1 + R = \bar{X}X,\tag{4}$$

This equation means that the autocorrelation of an earthquake seismograph is related to the reflection seismogram at the layer, or in other words: the correlation of ambient noise at the receiver, gives information about the channel. The 3D generalization was proved later on by different approaches, for example, by a power reciprocity theorem in [21], and (4) evolves into (5):

$$R(\vec{x}_A, \vec{x}_B, t) + R(\vec{x}_A, \vec{x}_B, -t) = 1 - T_{obs}(\vec{x}_A, -t) * T_{obs}(\vec{x}_B, t),\tag{5}$$

where $R(\vec{x}_A, \vec{x}_B, t)$ is the wavefield at a receiver located on $\vec{x}_A$ from a transmission happening at $\vec{x}_B$, and:

$$T_{obs}(\vec{x}_A, t) = \sum_i T(\vec{x}_A, \vec{x}_i, t) * N_i(t),\tag{6}$$

and analogously for $T_{obs}(\vec{x}_B, t)$. In (6), $T(\vec{x}_A, \vec{x}_i, t)$ is the response at a receiver on $\vec{x}_A$ from a distribution of N_i mutually uncorrelated noise sources located at $\vec{x}_i$ in the bottom half-space, and * is the convolution

operation. The meaning of the causal and acausal parts is that the response in both propagation directions between the pair of microphones at $\vec{x}_A$ and $\vec{x}_B$ is obtained, one for each sign of t.

The main conclusion from (5) is that by cross-correlating time-synchronized recordings of noise signals at two locations on the surface (right side of the equation), it is possible to reconstruct the signal that would be obtained if one of the locations was acting as transmitter, and the other one as receiver (left side). This means that a transmitter-receiver system can be replaced by a purely passive system based on the cross-correlation of acoustic noise.

This conclusion was also confirmed in parallel by [22], when performing experiments on acoustic thermal fluctuations. More specifically, they demonstrated that it is the time derivative of the cross-correlation function the signal that is equal to the transmitter-receiver system. Many works omit this time derivation as it can enhance undesirable noise. That is usually an acceptable approximation, unless accurate arrival times are needed [23].

As it was mentioned before, it was assumed that the medium was homogeneous and absorption-less. Later studies showed the validity of the approach without making assumptions for the medium, other than the presence of mutually uncorrelated noise sources, which leads to a diffuse acoustic field [24]. In realistic scenarios though, inhomogeneities in the medium, absorption, and the spatio-temporal distribution of realistic noise sources create a modified replica of the signal between the two receivers, where the amplitude of the response is particularly affected [23].

Finally, after the response between the two receivers is reconstructed from the noise cross-correlations, the first arrival can be discovered by a peak detector or a threshold operation, providing the range between the two sensors.

3. Experimental setup and signal processing steps

This section describes the setup of the experimental tests that were conducted to evaluate the feasibility of this technique for indoor environments such as an office room, as well as the required signal processing steps.

3.1. Description of the experiment

We conducted different tests in an office at the Department of Informatics, University of Oslo. The setup of the experiment can be seen in Fig. 2. We used two measurement microphones Behringer ECM8000, which have a flat response from 20 Hz to 20 kHz [25]. They were connected to a Zoom L-12 LiveTrack recorder [26], which synchronized the recordings from both microphones. Approximately 30 minutes of audio were stored in an SD card, with a resolution of 24 bits and a sampling frequency of 96 kHz. No channel gain was added. No one was present in the office during that time to avoid additional sound sources that could be created inadvertently, such as breathing or clothing rustle, that would deviate even more the ambient noise field from being diffuse. Fig. 2 also shows the real distance between the microphones, which was estimated manually with a tape measure. This ground truth distance had a value of 18.8 cm. The recorded data were processed offline in Matlab, following the steps described in the next subsection.
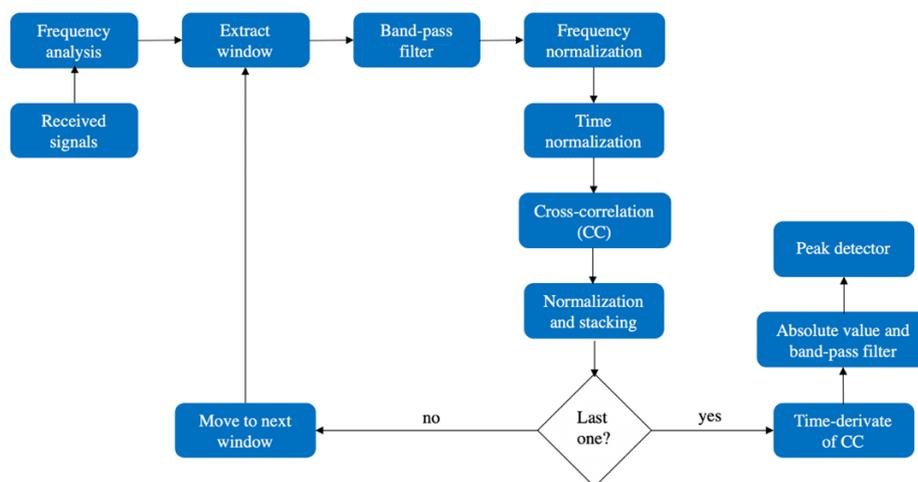
3.2. Signal processing steps

All the assumptions that were made to derive the theory, together with the typically low Signal-to-Noise Ratio (SNR) of ambient noise compared to active signals, imply that several processing steps are needed to recover the response between the two receivers. A good overview of different steps and their performance was given in [27]. In this section we describe the signal processing steps that we took in the experiment to obtain the ranges between the two microphones based on the cross-correlation of their recorded noise signals.

Figure 2: Set up of the experiment.

Fig. 3 shows the block diagram of the algorithm. First, after loading the received signals from both microphones, a frequency analysis is conducted. This analysis consists of an evaluation of the spectrogram and the coherogram between the two recorded signals. Fig. 4 shows the spectrogram from the first 10 minutes of data recorded at microphone 1 (on the left in Fig. 2). A similar spectrogram was obtained from microphone 2. This spectrogram was calculated from 1-second frames using a Hamming window, considering 50% overlap between frames, and frequencies between 20 Hz and 10 kHz. The strongest signals were obtained at the beginning of the recording, due to the noise created when leaving the office and locking the door. To avoid this interference, for the rest of the experiment the first two minutes of the signal were discarded. Other events with smaller amplitude can be observed later, probably caused by people passing by the office, and a nearby tram.


Figure 3: Block diagram of the signal processing algorithm.

The coherogram gives information about the better bandwidth (higher coherence) to process the recorded signals. In this case, it was calculated for the 30-minute signal by using one-minute frames with a window of 100 ms and 50% overlap, considering frequencies between 20 Hz and 10 kHz. These magnitude-squared coherence values were then averaged for all the frames. The result is shown as a blue line on Fig. 5, whereas the red line represents the moving mean of the averaged coherogram, which smooths the result.
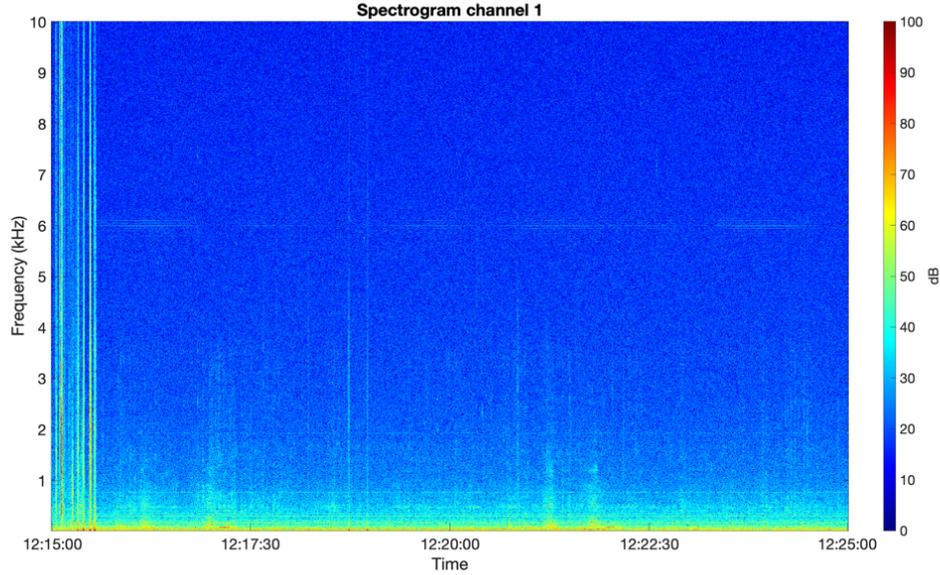
Figure 4: Spectrogram of the recorded signal at microphone 1.
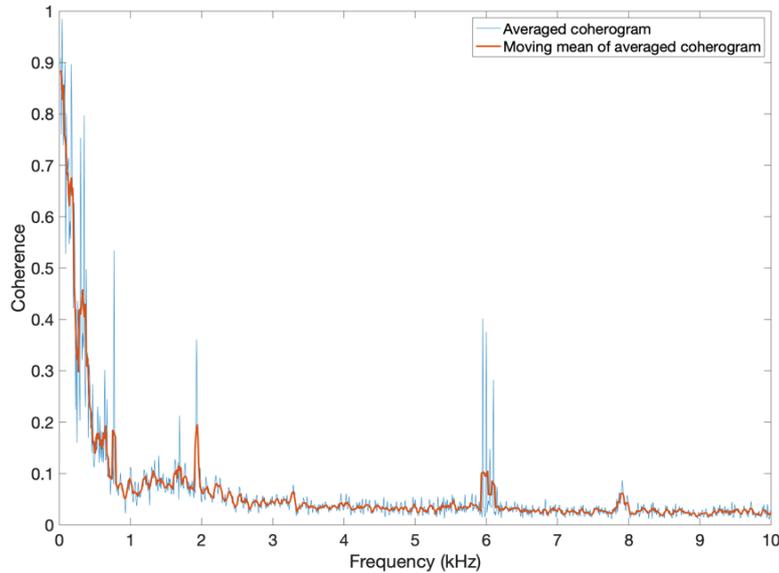
Figure 5: Averaged coherogram between the two recorded signals (blue) and moving mean of the averaged coherogram (red).

It can be observed from Fig. 5 that the frequencies with higher coherence values are below 200 Hz. There are also some smaller peaks around 2 and 6 kHz, caused by some tonal signals that can be observed in Fig. 4. Even if the highest coherence happens between 20 and 200 Hz, we had to include also higher frequencies, as temporal resolution is related to the inverse of the bandwidth [19]. In this case, assuming a nominal sound speed c of 343 m/s, the time-of-flight t_{TOF} between the microphones is:

$$t_{TOF} = \frac{d}{c} = \frac{0.188}{343} = 0.548 \text{ ms} \tag{7}$$

This value corresponds to a bandwidth $B \approx 1/t_{TOF} = 1824.5$ Hz, needed to resolve the travel time between the two microphones. Based on this result, we defined the bandwidth of interest between 40 and 2500 Hz, in order to ensure enough temporal resolution to detect the direct path.

Following the block diagram of Fig. 3, after the frequency analysis a window of a certain duration is extracted from both recorded signals. In this work we considered window frames of 100 ms, which are then band-pass filtered between 40 and 2500 Hz using a Kaiser window. One frame of the received signal for microphone 1 after band-pass filtering is shown in Fig. 6 (a).

In the next step frequency normalization is applied, which in this work consists of a spectrum whitening operation. This whitened spectrum is obtained by weighting the complex spectrum by a smoothed version of the amplitude spectrum [28]. This step compensates the higher attenuation experienced by higher frequency signals inside the selected bandwidth. Time normalization follows next, which applies a one-bit normalization to the signals, keeping only the sign, and not the amplitude. This step reduces the influence of strong discrete sources that deviate the noise field from being diffuse. The resulting signals are cross-correlated, and the correlation signal is normalized and accumulated. The normalized cross-correlation for one frame is given in Fig. 6 (b), where positive and negative distances refer to both propagation directions between the pair of microphones. It can be observed that there are no predominant peaks yet. The maximum is located outside the analysis window of 60 cm.
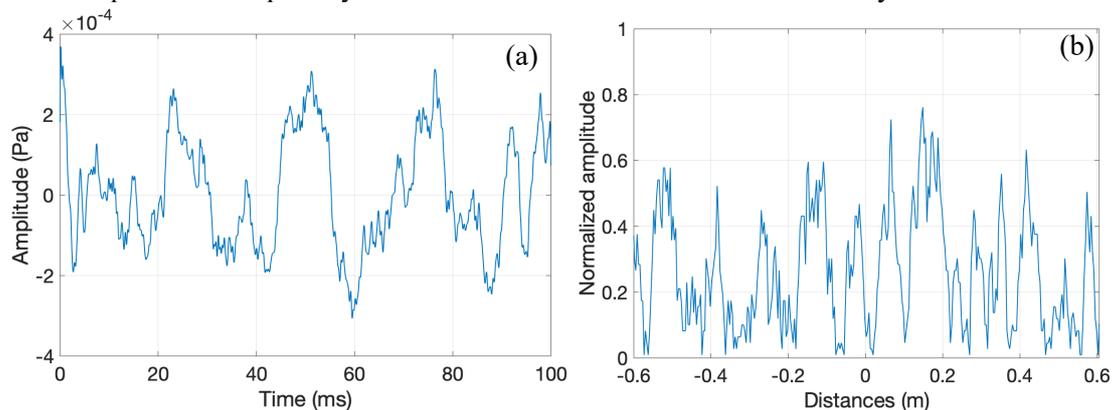

Figure 6: (a) 100 ms frame from the received signal at microphone 1, after bandpass-filtering; (b) absolute value of the cross-correlation from one frame.

If there are more frames to be evaluated, the algorithm moves to the next one considering an overlap of 25%, and it repeats the process until reaching the end of the file, or a pre-defined maximum evaluation time. The time derivative of the accumulated correlation signal is calculated then, and the resulting function is band-pass filtered and normalized. A peak detector can be applied at the end to find the first arrival in both directions, which indicates the range between the two microphones.

4. Experimental results

This section shows the obtained results for the ranging experiments conducted in the office. These results have been obtained by following the steps described in Fig. 3. As mentioned before, we skipped the first two minutes of data to avoid the strong noise produced when leaving the office.

First, we evaluated different accumulation times, in order to assess how much time is needed to recover the response between the microphones. Figure 7 shows the absolute value of the accumulated cross-correlations after band-pass filtering, for accumulation times of 500 ms (blue), 1 s (red) and 5 s (yellow). The same band-pass filter between 20 and 2500 Hz was used to smooth the cross-correlation. The ground truth is shown as dashed vertical black lines at the expected distance for both propagation directions (causal and acausal parts). The effect of the accumulation time is clearly noticeable now: with an accumulation time of 500 ms, the right peak is correctly recovered, but the left one still shows low amplitude. However, after accumulating 5 s both peaks are clearly distinguishable. This difference in the needed accumulation time for both peaks can be caused by the non-diffuse nature of the noise field. The sidelobes are also reduced after accumulating more correlations, as expected after increasing the SNR by the stacking process. For an accumulation time of 5 s, the peaks appeared at a distance of 18.04 cm (right peak), and 18.76 cm (left peak), obtaining a maximum error of 7.6 mm.
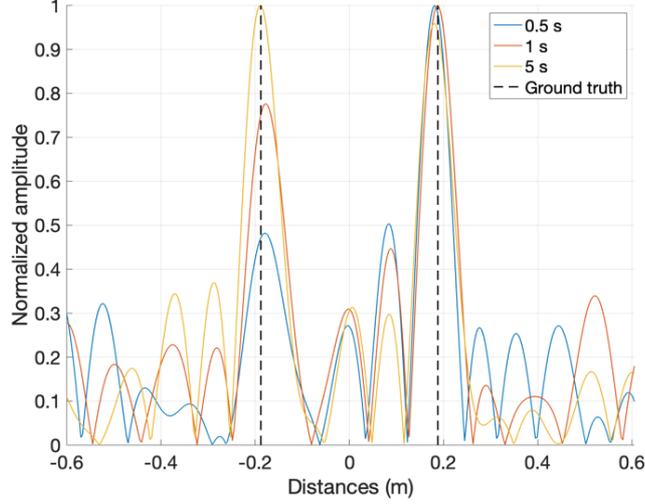
Figure 7: Absolute value of the accumulated cross-correlations after band-pass filtering, for different accumulation times.

We evaluated next the repeatability of the results by performing three additional experiments. In experiment 1, we calculated a new range estimate every minute for 10 minutes (10 measurements in total), considering an accumulation time of 5 seconds for each measurement. The mean value (μ) and standard deviation (σ) for the ranges were 18.90 ± 1.19 cm (left peak) and 17.54 ± 1.07 cm (right peak).

For experiment 2 we focused on a particularly quiet time window between 12:22:30 and 12:25:00 (see Fig. 4), where we could expect the noise to be more diffuse. We set again an accumulation time of 5 s and collected one new measurement every 10 s (10 measurements in total). Obtained range values were now 18.54 ± 1.39 cm (left peak) and 18.25 ± 1.60 cm (right peak). A slightly larger standard deviation than experiment 1 seems to indicate more variability in the estimations, although now the mean value obtained for the right peak is closer to the ground truth. The mean value for the left peak deviates from the ground truth, although not noticeably. This stronger variability in the results reflected in the standard deviations might indicate that more accumulation time is needed in quieter conditions. We then repeated the experiment in the same time window but accumulating now 10 s of cross-correlations, while taking one new measurement every 15 s (experiment 3). The obtained range values were 18.11 ± 1.24 cm (left peak) and 17.68 ± 0.86 cm (right peak). Both standard deviations are smaller in this case compared to the previous experiment, although the mean values deviate slightly from the ground truth. In any case, all the results are very close to the ground truth, showing that the method operates well under different conditions. The repeatability results have been gathered into Table 1, together with the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE).

Table 1
Experimental results obtained in the repeatability tests.

	Left peak	Right peak
Experiment 1		
$\mu \pm \sigma$ (cm)	18.90$\pm$1.19	17.54$\pm$1.07
MAE (cm)	0.93	1.26
RMSE (cm)	1.14	1.62
Experiment 2		
$\mu \pm \sigma$ (cm)	18.54$\pm$1.39	18.25$\pm$1.60
MAE (cm)	1.09	1.37
RMSE (cm)	1.34	1.61
Experiment 3		
$\mu \pm \sigma$ (cm)	18.11 $\pm$1.24	17.68$\pm$0.86
MAE (cm)	1.08	1.18
RMSE (cm)	1.36	1.38

5. Conclusions and future work

We have developed and evaluated in this work a ranging method based on the cross-correlation of ambient noise signals present in an indoor environment. This technique requires the signals from the microphones to be recorded synchronously, and they also need to be processed to increase the SNR and to reduce the effect of discrete noise sources, in order to recover the direct path between microphones. This could be solved by a communication link with a central server, either by a wire, or by Wi-Fi, which could gather the recorded signals and perform the processing.

Different tests were conducted in an office using two microphones separated 18.8 cm. The range was correctly calculated for both propagation directions after accumulating noise cross-correlations from 5 seconds of data. Different time windows from the recorded data were also evaluated, recovering the correct range in all cases with a minimum and maximum Mean Absolute Error of 0.93 and 1.37 cm.

These results show the feasibility of this technique for obtaining ranges between microphones in a purely passive way in an indoor environment. Future work will evaluate this technique on a more complex and realistic scenario, with several microphones deployed at larger distances between each other. This ranging method could be used for the passive self-calibration of beacons for indoor positioning applications, or nodes in a sensor network.

6. Acknowledgements

7. References

[1]   D. Estrin, L. Girod, G. Pottie and M. Srivastava, Instrumenting the world with wireless sensor networks, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, USA, 2001, pp. 2033–2036. doi: 10.1109/ICASSP.2001.940390.

[2]   D.E. Manolakis. "Efficient solution and performance analysis of 3-D position estimation by trilateration." IEEE Transactions on Aerospace and Electronic Systems 32.4 (1996): 1239–1248.

[3]   A. Mahajan and F. Figueroa. "An automatic self-installation and calibration method for a 3D position sensing system using ultrasonics." Robotics and Autonomous Systems 28.4 (1999): 281–294.

[4]   J. Ureña, D. Ruiz, J.C. García, J.J. García, A. Hernández and M.C. Pérez, LPS self-calibration method using a mobile robot, in: Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, Hangzhou, China, 2011, pp. 1–6. doi: 10.1109/IMTC.2011.5944284.

[5]   J. Wendeberg, F. Höflinger, C. Schindelhauer and L. Reindl, Anchor-free TDOA Self-Localization, in: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation, Guimarães, Portugal, 2011, pp. 1–10. doi: 10.1109/IPIN.2011.6071909

[6]   S. Thrun. "Affine structure from sound." Advances in Neural Information Processing Systems 18 (2006): 1353–1360.

[7]   J. Bordoy, C. Schindelhauer, F. Höflinger and L.M. Reindl. "Exploiting Acoustic Echoes for Smartphone Localization and Microphone Self-Calibration." IEEE Transactions on Instrumentation and Measurement 69.4 (2020): 1484–1492.

[8]   P.D. Jager, M. Trinkle, and A. Hashemi-Sakhtsari, Automatic Microphone Array Position Calibration Using an Acoustic Sounding Source, in: Proceedings of the IEEE Conference on Industrial Electronics and Applications, Xi'an, China, 2009, pp. 2110–2113. doi: 10.1109/ICIEA.2009.5138521

[9]   A. Savvides, C-C. Han and M.B. Strivastava, Dynamic fine-grained localization in ad-hoc networks of sensors, in: Proceedings of the Annual International Conference on Mobile Computing and Networking, Rome, Italy, 2001, pp. 166–179. doi: 10.1145/381677.381693.

[10] V.C. Raykar and R. Duraiswami, Automatic Position Calibration of Multiple Microphones, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 2004, pp. 69–72. doi: 10.1109/ICASSP.2004.1326765.

[11] L. Girod, M. Lukac, V. Trifa and D. Estrin, The Design and Implementation of a Self-Calibrating Distributed Acoustic Sensing Platform, in: Proceedings of the International Conference on Embedded Networked Sensor Systems, Boulder, CO, USA, 2006, pp. 71–84. doi: 10.1145/1182807.1182815.

[12] C. De Marziani, J. Ureña, A. Hernández, M. Mazo, J.J. García, A. Jiménez, M.C. Pérez, F.J. Álvarez and J.M. Villadangos. "Acoustic sensor network for relative positioning of nodes." Sensors 9 (2009): 8490–8507.

[13] I. McCowan, M. Lincoln and I. Himawan. "Microphone Array Shape Calibration in Diffuse Noise Fields." IEEE Transactions on Audio, Speech, and Language Processing 16.3 (2008): 666–670.

[14] M.J. Taghizadeh, A. Asaei, P.N. Garner and H. Bourlard, Ad-hoc microphone array calibration from partial distance measurements, in: Proceedings of the Joint Workshop on Hands-Free Speech Communication and Microphone Arrays, Villers-lès-Nancy, France, 2014, pp. 1–5. doi: 10.1109/HSCMA.2014.6843239.

[15] J. Velasco, M.J. Taghizadeh, A. Asaei, H. Bourlard, C.J. Martín-Arguedas, J. Macias-Guarasa and D. Pizarro, Novel GCC-PHAT model in diffuse sound field for microphone array pairwise distance based calibration, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, South Brisbane, QLD, Australia, 2015, pp. 2669–2673. doi: 10.1109/ICASSP.2015.7178455

[16] J.F. Claerbout. "Synthesis of a layered medium from its acoustic transmission response." Geophysics 33.2 (1968): 264–269.

[17] N.M. Shapiro, M. Campillo, L. Stehly and M.H. Ritzwoller. "High-resolution surface-wave tomography from ambient seismic noise." Science 307.5715 (2005): 1615–1618.

[18] K.G. Sabra, P. Roux, P. Gerstoft, W.A. Kuperman and M.C. Fehler. "Extracting coherent coda arrivals from cross-correlations of long period seismic waves during the Mount St. Helens 2004 eruption." Geophysical Research Letters 33.6 L06313 (2006): 1–4.

[19] K.G. Sabra, P. Roux, A.M. Thode, G.L. D'Spain, W.S. Hodgkiss and W.A. Kuperman. "Using ocean ambient noise for array self-localization and self-synchronization." IEEE Journal of Oceanic Engineering 30.2 (2005): 338–347.

[20] J.F. Claerbout, Acoustic daylight imaging: Introduction to the underlying concept: A prospect for the instrumented oil field, Technical Report SEP-108, Stanford University, Stanford, CA, United States, 2001.

[21] K. Wapenaar. "Synthesis of an inhomogeneous medium from its acoustic transmission response." Geophysics 68.5 (2003): 1756–1759.

[22] R.L. Weaver and O.I. Lobkis. "Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies." Physical Review Letters 87.13 ID 134301 (2001): 1–4.

[23] P. Roux, K.G. Sabra and W.A. Kuperman. "Ambient noise cross correlation in free space: Theoretical approach." The Journal of the Acoustical Society of America 117.1 (2005): 79–84.

[24] K. Wapenaar. "Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation." Physical Review Letters 93.25 ID 254301 (2004): 1–4.

[25] Behringer, ECM8000, 2021. URL: https://www.behringer.com/product.html?modelCode=P0118

[26] Zoom, LiveTrack L-12, 2021. URL: https://zoomcorp.com/en/ca/digital-mixer-multi-track-recorders/digital-mixer-recorder/livetrak-l-12/

[27] L.A. Brooks and P. Gerstoft. "Green's function approximation from cross-correlations of 20-100 Hz noise during a tropical storm." The Journal of the Acoustical Society of America 125.2 (2009): 723–734.

[28] G.D. Bensen, M.H. Ritzwoller, M.P. Barmin, A.L. Levshin, F. Lin, M.P. Moschetti, N.M. Shapiro and Y. Yang. "Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements." Geophysical Journal International 169.3 (2007): 1239–1260.