Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

Self-Localization of Single Microphone from Ultrasound Field Generated by Sources with Subwavelength Displacement

Keisuke Hasegawa, Takaaki Nara
Graduate School of Information Science and technology
the University of Tokyo
Tokyo, Japan
keisuke_hasegawa@ipc.i.u-tokyo.ac.jp, nara@alab.t.u-tokyo.ac.jp

***Abstract*—We propose a self-localization method of a single microphone based on measured phase gradient of wave field created by sources with sub-wavelength movements. Unlike most of conventional localization methods based on time-of-arrival (TOA) of time-difference-of-arrival (TDOA) from impulsive wave packets emitted from sources, we employ spatial information of wave field that is entirely derived from the wave equation containing a point source. Instead of impulsive emissions, continuous sinusoidal wave from a point source is utilized in our technique. We describe a localization strategy that integrates multiple measurements for a stable and robust performance. We verified the validity of our method via numerical and real-environment experiments.**

***Index Terms*—Self-Localization, Single Microphone, Phase Gradient Field**

I. Introduction

Recently, small-sized autonomous devices such as smart phones, quadcopters, wearable devices, and so forth, are frequently used in our daily lives. They are installed with functions such as measurement, communication and information display. Self-localization with small sensors is one of the most fundamental abilities that they desire to possess for its quite wide applications.

However, there has been no such a definitive method other than GPS (Global Positioning System), which gives meter-order estimation in coarse temporal resolution only valid for outdoor use. This situation is a contrast to self-posture estimation that is accomplished by inertial measurement unit (IMU) sensor [1] that is nowadays as small as a cubic of several millimeters. Video-based techniques are one of promising methodologies [2]. However, they require prior learning phase for capturing environmental information.

In acoustics, there has been several attempts to realize a self-localization of a single microphone out of received signals with prior knowledge about the sound emission. One example is the use of reverberant information of the room [3] and another is the location of the sources [4]. With those methods, it is common to use the signal generation model based on the impulse response function or TOA/TDOA based localization technique of received pulses.

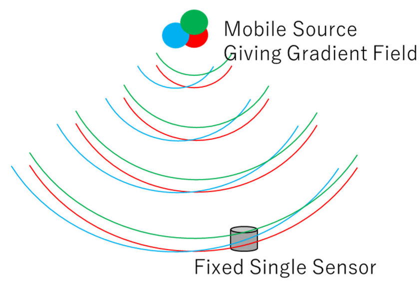

Fig. 1. Our strategy to achieve single microphone localization based on sources with small tunable displacement.

Here we propose a new method for self-localization that directly refers to spatial information of sound field entirely described by the wave equation. We postulate the use of point sources whose spatial displacement can be tuned by the degrees smaller than the wavelength (Fig. 1). Instead of pulse wave, the sources are supposed to emit continuous sinusoidal sound wave. We utilize the phase information of received signals by processing it with the quadrature detection technique. Thanks to this, the proposed method is expected to be robust against non-stationary impulsive environmental noises.

II. Basic Equations

In general, pressure field $p(\boldsymbol{r}, t)$ follows the wave equation:

$$\left(\nabla^2 + \frac{\partial^2}{\partial t^2}\right) p(\boldsymbol{r}, t) = 0, \quad (1)$$

where $\boldsymbol{r} := [x\ y\ z]^T$ denotes three-dimensional location in space and t denotes time, respectively. Hereinafter let $\boldsymbol{r}$ be a position of the microphone. Suppose that we have a single point source with no directivity at $\boldsymbol{r}_0$. It is known that for this situation the following equation is obtained by expressing (1) in the polar coordinates system [5]:

$$\left(|\boldsymbol{r} - \boldsymbol{r}_0|\nabla_{\boldsymbol{r}} + \frac{\boldsymbol{r} - \boldsymbol{r}_0}{|\boldsymbol{r} - \boldsymbol{r}_0|} + \frac{\boldsymbol{r} - \boldsymbol{r}_0}{c}\frac{\partial}{\partial t}\right) p(\boldsymbol{r}, t) = \boldsymbol{0}, \quad (2)$$

where c is the sound velocity and $\nabla_{\boldsymbol{r}}$ indicates the gradient operator for $\boldsymbol{r}$. Since (2) is a vector expression of a system of three equations, it is in principle possible that the source position $\boldsymbol{r}_0$ can promptly be calculated if the microphone position $\boldsymbol{r}$, the instantaneous pressure $p(\boldsymbol{r}, t)$ and its spatial gradient $\nabla_{\boldsymbol{r}} p(\boldsymbol{r}, t)$ are all obtained.

In (2), the pressure field is a function of both the receiving point $\boldsymbol{r}$ and the source position $\boldsymbol{r}_0$. For this reason it should be rewritten as $p(\boldsymbol{r}, \boldsymbol{r}_0, t)$. Since we assume to have a point source in the field, it is evident that the pressure field is entirely dependent on the distance from the source $|\boldsymbol{r} - \boldsymbol{r}_0|$ and t. Therefore, the following expression holds with a function $A(\boldsymbol{r}, t)$:

$$p(\boldsymbol{r}, \boldsymbol{r}_0, t) = A(|\boldsymbol{r} - \boldsymbol{r}_0|, t). \tag{3}$$

Here we define another gradient operator $\nabla \boldsymbol{r}_0$, denoting spatial differentiation about $\boldsymbol{r}_0$. For (3), the following relation is obtained:

$$\nabla_{\boldsymbol{r}_0} A(|\boldsymbol{r} - \boldsymbol{r}_0|, t) = -\nabla_{\boldsymbol{r}} A(|\boldsymbol{r} - \boldsymbol{r}_0|, t). \tag{4}$$

With this equation, (2) can be rewritten as a form where source position $\boldsymbol{r}_0$ and microphone position $\boldsymbol{r}$ is exchanged:

$$\left(|\boldsymbol{r}_0 - \boldsymbol{r}| \nabla_{\boldsymbol{r}_0} + \frac{\boldsymbol{r}_0 - \boldsymbol{r}}{|\boldsymbol{r}_0 - \boldsymbol{r}|} + \frac{\boldsymbol{r}_0 - \boldsymbol{r}}{c} \frac{\partial}{\partial t} \right) p(\boldsymbol{r}, \boldsymbol{r}_0, t) = \boldsymbol{0}. \tag{5}$$

This means that the microphone location $\boldsymbol{r}$ at a moment can be known if concurrent instantaneous pressure p and its spatial gradient about the source position $\nabla_{\boldsymbol{r}_0} p$ are measured by the microphone and the source position $\boldsymbol{r}_0$ is known. $\nabla_{\boldsymbol{r}_0} p$ is understood as the ratio of change in the observed pressure waveform to a slight spatial fluctuation of the source position.

III. Localization Strategy

It is expected that based on (5), we can directly determine the microphone position with a single-shot measurement from a source. Although this is theoretically correct, such a strategy contains a couple of problems. Simultaneous measurement of an instantaneous pressure and its three-dimensional gradient at a certain point requires a particularly designed microphone. In addition, a single instantaneous measurement is quite vulnerable to noises. For these reasons, we utilize integrated multiple measurements in order to a stable and robust localization.

We assume that the source position can be tuned by a sub-wavelength span. For realizing this, use of mechanical structure such as ultrasonic-motor-based tables is one solution. Constructing an array of multiple tiny loudspeakers in an spatially integrated form as seen in [6] is an alternative method. Use of a phased array [7] [8] based on sound field synthesis [9] for generating a virtual point source with tunable location is another effective way. The spatial pressure gradient $\nabla_{\boldsymbol{r}_0} p|_{\boldsymbol{r}_0}$ is then given as a differentially approximated form:

$$\nabla_{\boldsymbol{r}_0} p|_{\boldsymbol{r}_0 = [x_0 \ y_0 \ z_0]^T} \approx \begin{bmatrix} \frac{p(x_0+d/2, y_0, z_0) - p(x_0-d/2, y_0, z_0)}{d} \\ \frac{p(x_0, y_0+d/2, z_0) - p(x_0, y_0-d/2, z_0)}{d} \\ \frac{p(x_0, y_0, z_0+d/2) - p(x_0, y_0, z_0-d/2)}{d} \end{bmatrix}, \tag{6}$$

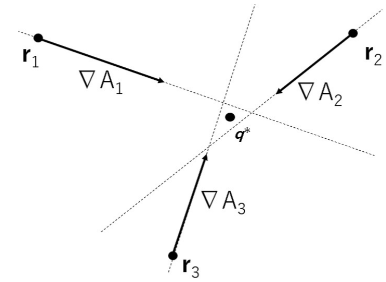

Fig. 2. Lines parallel to the gradient ideally converges at the microphone location.

Here d denotes the source displacement smaller than the wavelength.

This means that multiple (typically 6 times) synchronized measurements is necessary for determining $\nabla_{\boldsymbol{r}_0} p|_{\boldsymbol{r}_0}$. In order to virtually achieve this, we assume that the source emits sinusoidal spherical wave with its frequency ω and that we can refer to measured complex pressure amplitude $P(\boldsymbol{r}, \boldsymbol{r}_0)$, meaning sequential observation that lasts for more than one waveform cycle ($T = 2\pi/\omega$) is presupposed. Instantaneous pressure values at an arbitrary time t is reconstructed from corresponding sinusoidal complex pressure amplitude:

$$p(\boldsymbol{r}, \boldsymbol{r}_0, t) = \mathrm{Re}[P(\boldsymbol{r}, \boldsymbol{r}_0) e^{j\omega t}], \tag{7}$$

where Re[$\cdot$] indicated the real part of $\cdot$. Hence adding source displacement by $\Delta \boldsymbol{r}_0$ and measuring $P(\boldsymbol{r}, \boldsymbol{r}_0 + \Delta \boldsymbol{r}_0, t))$ for several $\Delta \boldsymbol{r}_0$ gives $\nabla_{\boldsymbol{r}_0} p(\boldsymbol{r}, \boldsymbol{r}_0, t)$ in (5) as an approximated form as shown in (6).

In the following, we assume that $\boldsymbol{r}_0$, $\nabla_{\boldsymbol{r}_0} p(\boldsymbol{r}, \boldsymbol{r}_0, t_0)$, $p(\boldsymbol{r}, \boldsymbol{r}_0, t_0)$, and $\frac{\partial p}{\partial t}(\boldsymbol{r}, \boldsymbol{r}_0, t_0)$ are all obtained for $t = t_0$. As stated, we do not directly solve (5), but consider integrating multiple measurements. The left-hand side of (5) contains the following vectors: $\nabla_{\boldsymbol{r}_0} p(\boldsymbol{r}, \boldsymbol{r}_0, t_0)$ and $(\boldsymbol{r}_0 - \boldsymbol{r})/|\boldsymbol{r}_0 - \boldsymbol{r}|$. Since they must cancel to vanish those two vectors must be parallel to each other, meaning that the source direction from the microphone location is given by the pressure gradient about the source position. Therefore, for multiple sources with different locations that give their inherent pressure gradients, a set of corresponding lines that passes sources and expands along the gradient is expected to cross at a specific point, which namely is supposed to coincide with the microphone location (figure 2). When it is possible to perform sequential measurements with multiple sources driven one by one, this strategy can be applied to microphone localization in an expectation of gradual decrease in estimation error as the number of sources increases. Note that this strategy uses no absolute amplitude value, which shows little spatial difference as the sources is located farther.

A reasonable criteria of localization would be a minimization of mean squared error from those lines corresponding to each source. Each line is expressed by a vector $\boldsymbol{q}$ and a scalar parameter α that follows the following equation:

$$\boldsymbol{q} = \boldsymbol{r}_0 + \alpha \nabla_{\boldsymbol{r}_0} p(\boldsymbol{r}, \boldsymbol{r}_0, t_0). \tag{8}$$

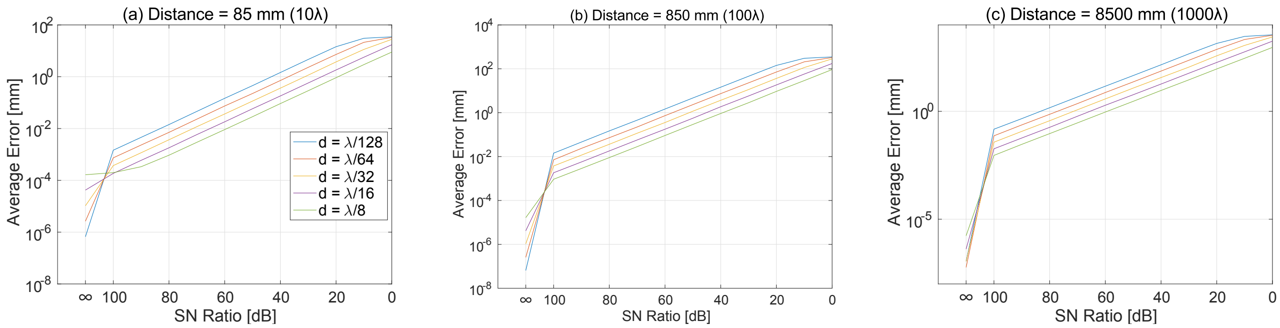

Fig. 3. Localization errors with the source-microphone distance of (a)10λ,(b)100λ, (c)1000λ.

The squared distance l^2 between this line and an external point $\boldsymbol{q}_0$ is given as

$$l^2 = \frac{|\boldsymbol{g}(\boldsymbol{r}_0)|^2|\boldsymbol{q}_0 - \boldsymbol{r}_0|^2 - (\boldsymbol{q}_0 - \boldsymbol{r}_0)^T\boldsymbol{g}(\boldsymbol{r}_0)\boldsymbol{g}(\boldsymbol{r}_0)^T(\boldsymbol{q}_0 - \boldsymbol{r}_0)}{|\boldsymbol{g}(\boldsymbol{r}_0)|^2}, \tag{9}$$

where $\boldsymbol{g}(\boldsymbol{r}_0) = \nabla_{\boldsymbol{r}_0} p(\boldsymbol{r}, \boldsymbol{r}_0, t_0)$. Let N be the number of sources, $\boldsymbol{r}_i, i = 1, \ldots, N$ be the source locations and $\boldsymbol{g}_i = \nabla_{\boldsymbol{r}_i} p(\boldsymbol{r}, \boldsymbol{r}_i, t_0)$ be their spatial gradient about $\boldsymbol{r}_0$. The location $\boldsymbol{q}^*$ giving the least mean squared distance described in (9) is obtained by solving the following linear system of equation:

$$\left(\sum_{i=1}^{N} \frac{|\boldsymbol{g}_i|^2 I - \boldsymbol{g}_i\boldsymbol{g}_i^T}{|\boldsymbol{g}_i|^2}\right)\boldsymbol{q}^* = \sum_{i=1}^{N} \frac{|\boldsymbol{g}_i|^2 I - \boldsymbol{g}_i\boldsymbol{g}_i^T}{|\boldsymbol{g}_i|^2}\boldsymbol{r}_i, \tag{10}$$

where I is a 3×3 unit matrix. The eventually obtained $\boldsymbol{q}^*$ is the estimated microphone location.

This strategy is highly dependent on the accuracy of the field gradient $\nabla_{\boldsymbol{r}_0} p(\boldsymbol{r},\boldsymbol{r}_0, t_0)$. It is equivalent to the accuracy of the complex amplitude $P(\boldsymbol{r},\boldsymbol{r}_0, t_0)$ since its spatial difference is what is given as gradient. The complex amplitude can be precisely estimated by quadrature detection method. However, obtaining instantaneous pressure by (7) occasionally magnify the contained errors due to measurement noises when the S/N ratio is low or the reconstructed pressure is close to the zero-cross point of the waveform. For a spherical source, the phase gradient and instantaneous pressure gradient at a certain location is in the same direction since it is also a function of $|\boldsymbol{r} - \boldsymbol{r}_0|$. Since the estimation of the phase $\angle P(\boldsymbol{r},\boldsymbol{r}_0, t_0)$ can be stable with sufficiently long measurement period, use of phase gradient instead of direct pressure gradient is expected to give a more stable estimation. In the following experiment, we used the phase gradient $\nabla_{\boldsymbol{r}_i} \angle P(\boldsymbol{r},\boldsymbol{r}_i, t_0)$ that replaced $\boldsymbol{g}_i = \nabla_{\boldsymbol{r}_0} p(\boldsymbol{r},\boldsymbol{r}_i, t_0)$ in (10).

IV. Experiments

A. Numerical Experiments

We first performed numerical simulations for assessing the validity of our proposed localization technique. We supposed that the received signal without noise is given as

$$p(\boldsymbol{r}, \boldsymbol{r}_0, t) = p_0 \frac{\cos(\omega t - k|\boldsymbol{r} - \boldsymbol{r}_0|)}{|\boldsymbol{r} - \boldsymbol{r}_0|}, \tag{11}$$

where $k = \omega/c$ is the wavenumber. We set the parameters as p_0 = 1Pa, ω = 40kHz, c = 340m/s. The phase gradient is calculated based on (6) after quadrature detection process to the waveform described in (11).

Factors that might affect the localization performance are the S/N ratio of received signal, source displacement d for approximating gradient and the distance $|\boldsymbol{r} - \boldsymbol{r}_0|$ between sources and microphones. In the first experiment, white noises are added to adjust the S/N ratio (-∞ to 0db) to the desired value. The source locations were randomly determined so that their distance from the microphone is consistent (10λ, 100λ and 1000λ with λ denoting the wavelength, approximately 8.5mm in the experiment.) We also varied the source displacement d among λ/8, λ/16, λ/32, λ/64 and λ/128. The number of sources were fixed to eight.

Figure 3 shows the mean localization errors [m] calculated from 1,000 independent measurements for each set of parameters. Overall, it is safely said that the localization containing mean error less than one millimeter was successfully done with a good S/N ratio. The result shows that the mean error increased as the source-microphone distance increased and the S/N ratio was lowered. It is observed that increase by 10 times in source-microphone distance had as much effect on error increase as S/N ratio deterioration by 20dB. In the absence of noises, smaller d gave better estimation. Conversely, larger d yielded smaller errors when the received signals contained noises even if they are quite small.

Next, we performed the same experiment with three sources. We found that the mean error did not significantly changed compared with that with eight sources. This is presumably because in this experiment the sources constructed a sparse but large aperture that surrounded the microphone.

B. Measurement-Based Experiments

We conducted the localization from recorded signals in the real environment. Figure 4 shows the experimental setup. As a source, we used a 10mm-diameter 40kHz ultrasonic transducer (MA40S4S, Murata Electronics). The signals were captured by a 3mm-diameter ECM via an digital oscilloscope (Picoscope 4262) and stored to PC. We performed an offline quadrature detection with an externally supplied sinusoidal signal which is supposed to be generated with an internal clock in the

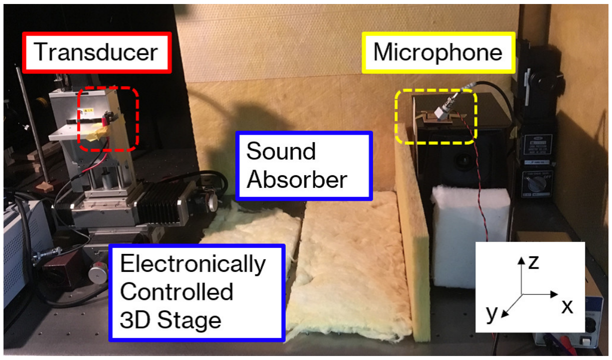

Fig. 4. Experimental setup.

practical situation. We determined two different microphone locations and a set of three sources. The source location were $(x, y, z) = (24, -24, -24), (-24, 24, -24), (-24, -24, 24)$ [mm]. For each source location, we performed six consecutive measurements in line with (11) in order to calculate the corresponding spatial phase gradient. The source displacement d was set to 2mm.

Figure 5 and Table 1 shows the experimental results. For the microphone located at $(165, 0, 10)$[mm], the gradient lines are seen to well converge around the estimated microphone location. On the other hand, with a microphone location $(535, 0, 10)$[mm], the lines are closer to be parallel to one another. This might cause estimation error especially in the depth direction. Note that the microphone location was configured manually and may contain several millimeter errors. Therefore, the estimation error in Table 1 is not something rigorous. It would rather be safer to say that the experimental results just indicates the rough verification of the proposed localization method being effective in some situations.

V. Conclusion

We have proposed a self-localization technique of a single microphone based on measurements of phase gradient from individual emission of multiple sources with displacement smaller than the wavelength. We have validated the efficacy of the proposed localization algorithm in a numerical experiment, in which we also quantitatively assessed the influence of systematic factors on estimation errors such as S/N ratio, source-microphone distances, source displacements and source numbers. In real-environment experiment, we demonstrated that our localization strategy appropriately worked and localization errors were several tens of millimeter.

In practical operation, unidirectional communication protocol on ultrasonic modulation must be determined for transmitting source location and displacement to the receivers. In addition, the phase-zero timing of the source must not fluctuate over the entire operation period. Our next challenges include construction of a real-time estimation system by fabricating sources with tunable displacements.

TABLE I
Estimation errors.

True Position [mm]	Error [mm]
(165, 0, 10)	(-31.7, 0.5, -1.2)
(535, 0, 10)	(-6.9, 4.2, 26.7)

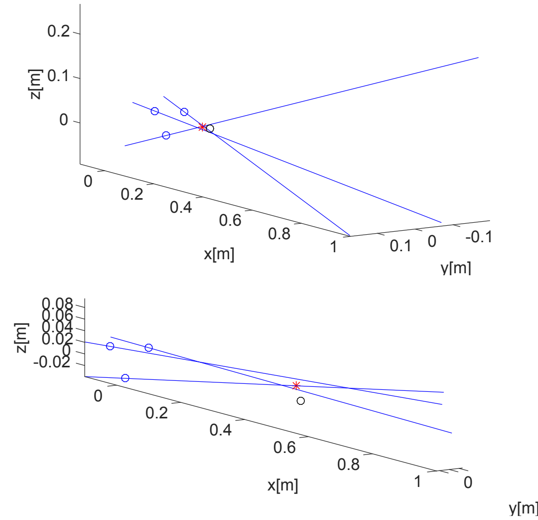

Fig. 5. Localization results and gradient lines for microphone located at (160, 0, 10) [mm] (Upper Figure) and at (535, 0, 10) [mm] (Lower Figure). Blue circles, black circles and red stars indicate the source positions, the true microphone positions, and the estimated microphone positions.

References

[1] S. O. H. Madgwick, A. J. L. Harrison and R. Vaidyanathan, “Estimation of IMU and MARG orientation using a gradient descent algorithm,” 2011 IEEE International Conference on Rehabilitation Robotics, Zurich, 2011, pp. 1-7.
[2] C. Cadena et al., ”Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age,” in IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309-1332, Dec. 2016.
[3] R. Parhizkar, I. Dokmanic and M. Vetterli, "Single-channel indoor microphone localization," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, 2014, pp. 1434-1438.
[4] D. B. Haddad, W. A. Martins, M. d. V. M. da Costa, L. W. P. Biscainho, L. O. Nunes and B. Lee, ”Robust Acoustic Self-Localization of Mobile Devices,” in IEEE Transactions on Mobile Computing, vol. 15, no. 4, pp. 982-995, 1 April 2016.
[5] S. Ando, T. Nara, and T. Levy, “Partial differential equation-based localization of a monopole source from a circular array,” J. Acoust. Soc. Am., vol. 134, no. 4, pp. 2799-2813, 2013.
[6] Nobutaka Ono, Nobutaka Ito and Shigeki Sagayama, ”Five Classes of Crystal Arrays for Blind Decorrelation of Diffuse Noise,” Proc. SAM, pp. 151-154, Jul. 2008.
[7] L. W. Schmerr, Jr., “Fundamentals of Ultrasonic Phased Arrays,” Springer International Publishing, 2015.
[8] Keisuke Hasegawa, Hiroyuki Shinoda, ”Aerial Vibrotactile Display Based on MultiUnit Ultrasound Phased Array,” IEEE Transactions on Haptics, vol. 11, no. 3, pp. 367-377, 1 July-Sept. 2018.
[9] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” J. Acoust. Soc. Am., vol. 93, no. 5, pp. 2764-2778, 1993.