Locating an Acoustic Source Using Low-Cost Hardware

Melanie Brinkschulte[1]

Abstract: The availability of cheap but powerful sensors and signal processing hardware allows the solution of problems, which previously could only be solved at high costs and efforts. In this paper we describe an approach to locate an acoustic sound source by three low cost microphones and operational amplifiers in combination with a simple microcontroller using its embedded analog/digital converters. There are several applications, e.g. the detection of the direction and distance of a police car by the sound of its siren or the location of a speaking person by an assistance robot in a household. The complete real-time analog/digital signal processing chain using cheap hardware is in the focus of this approach. By deriving an approximate solution for the location equations, the approach becomes suitable for low-cost microcontrollers.

Keywords: Sound Source; Location; Low-Cost Hardware

1 Introduction

The availability of cheap but powerful sensors and signal processing hardware allows the solution of problems, which previously could only be solved at high costs and efforts. Therefore, research in the field of cyber-physical systems has also largely benefited from this progress in hardware. To combine and fuse sensor data with limited hardware resources and to process this data with sufficient accuracy is a major challenge. In this paper we describe an approach to locate an acoustic sound source by three low cost microphones and operational amplifiers in combination with a simple microcontroller using its embedded analog/digital converters. There are several applications, e.g. the detection of the direction and distance of a police car by the sound of its siren or the location of a speaking person by an assistance robot in a household. The complete real-time analog/digital signal processing chain using cheap hardware is in the focus of this approach.

The paper is structured as follows: After the introduction and motivation, the basic ideas and conceptions are presented in section 2. Section 3 describes the overall architecture and the processing software of our approach. Section 4 presents the evaluation results. In section 5, related work is described while section 6 concludes this paper.

[1] Goethe University Frankfurt am Main, Computer Science Department, Robert-Maier-Str. 11-15, 60325 Frankfurt am Main, Germany MelanieBrinkschulte@t-online.de

doi:10.18420/in2017_89

2 Basic Conception

The basic idea to localize an acoustic sound source is to measure the phase shift between microphones placed at different locations, as shown in Figure 1. The measured phase shift can be transformed to the distance difference Δl between the microphones according to the following formula:

$$\Delta l = \tau \cdot v_{Sound} \tag{1}$$

To allow an exact unambiguous localization, the minimum microphone configuration is shown in Figure 2. To determine the flat coordinates (x, y) of the sound source location in a plane, at least two measured distance differences Δl_1 and Δl_2 are needed for a unique solution of the resulting equation system. Thus, three microphones are necessary. Furthermore, the three microphone should not be placed along a single axis because the mirrored solutions in front and back of this axis could not be distinguished. According to Figure 2, the following equations have to be solved to determine the position of the sound source:

$$x^2 + y^2 = l^2 \qquad (x+d)^2 + y^2 = (l + \Delta l_1)^2 \qquad (x+d)^2 + (y+d)^2 = (l + \Delta l_2)^2$$

Unfortunately, the exact solution of this equation system is rather complex and not well

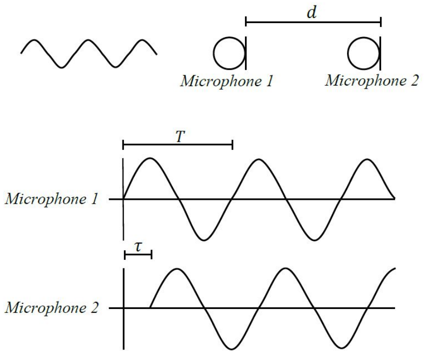

Fig. 1: Phase shift as the basic principle of distance measurement

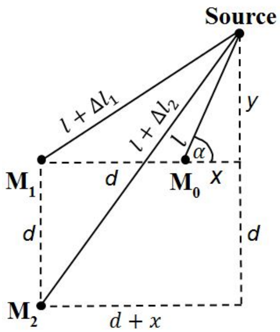

Fig. 2: Position of three microphones

suited for the target platform, a small and cheap microcontroller (see [Br16] for the exact solution). Therefore, a less complex approximation has been found. Figure 3 sketches how to approximate the angle of the sound source in relation to two microphones. For larger distances, the following approximations $\overline{M_0X} \approx \overline{H_0X}$ and $\overline{M_1X} \approx \overline{H_1X}$ hold true. Using Trigonometry, the following equations can be derived:

$$\overline{H_0Z} = \overline{H_1Z} = \frac{d}{2} \cdot \cos(\alpha) \Rightarrow \Delta l \approx \overline{H_1X} - \overline{H_0X} = d \cdot \cos(\alpha) \Rightarrow \alpha \approx \arccos\left(\frac{\Delta l}{d}\right) \tag{2}$$

To decide the quality of this approximation, we can compare the Δl given by (2) with the exact solution[2] for Δl derived from Figure 3:

[2] It is easy to exactly calculate Δl from α; however, the reverse calculation of α from Δl is rather complex

$$\Delta l = \sqrt{\left(l \cdot cos(\alpha) + \frac{d}{2}\right)^2 + (l \cdot sin(\alpha))^2} - \sqrt{\left(l \cdot cos(\alpha) - \frac{d}{2}\right)^2 + (l \cdot sin(\alpha))^2}$$

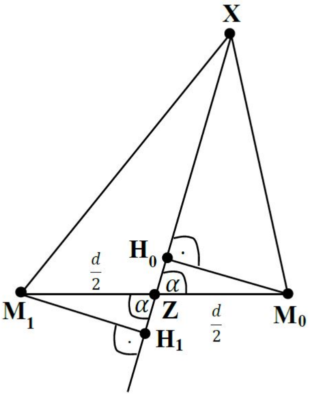

Fig. 3: Angle approximation using 2 microphones

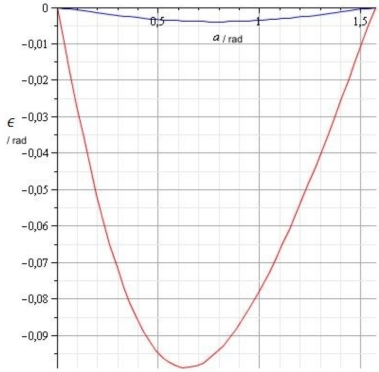

Fig. 4: Angle imprecision caused by the approximation

Using this equation we can calculate the exact value for Δl for a given angle α and then calculate the approximated value for α using equation (2). The difference of the two angles gives the angle imprecision. Figure 4 shows this imprecision examplarily for two distances of $l = 1$ meter (blue curve) and $l = 0.2$ meter (red curve). The microphone distance d has been chosen to 0.245 meters[3]. It shows the angle imprecision is rather small and acceptable for larger distances. Even at the small distance of 0.2 meters the maximum angle imprecision is about 0.1 rad (5.7°).

The above approximation can be used to calculate the angle of the sound source. However, the distance cannot be calculated. Furthermore, there exists a second solution that is mirrored along the axis of the two microphones. Both issues can be solved using again the L-shaped 3 microphone configuration as shown in Figure 5. The distance can be calulated by triangulation using two approximated angles α_1 and α_2. This leads to the following equations[4]:

[3] This is the microphone distance of our experimental setup used for the evaluation

[4] The special cases ($\alpha_1 = 90°$, $\alpha'_2 = 90°$) produce simple solutions (($x = \frac{d}{2}$, $y = \tan(\alpha'_2) \cdot \frac{d}{2} - \frac{d}{2}$), ($x = 0$, $y = -\frac{d}{2} \cdot \tan(\alpha_1)$)). The special case $\alpha_1 = \alpha'_2 = 45°$ produces no solution.

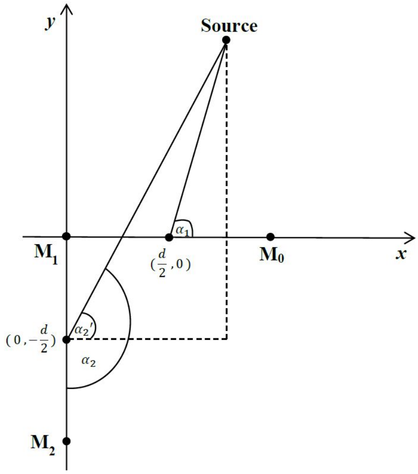

Fig. 5: Distance approximation using 3 microphones

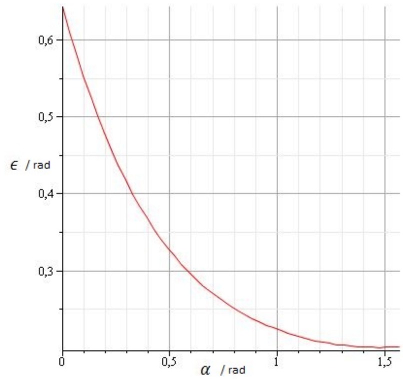

Fig. 6: Angle imprecision caused by a distance imprecision of 5 cm

$$y = \tan(\alpha_1) \cdot x - \tan(\alpha_1) \cdot \frac{d}{2} = \tan(\alpha_2') \cdot x - \frac{d}{2} \tag{3}$$

$$x = \frac{d}{2} \cdot \frac{\tan(\alpha_1) - 1}{\tan(\alpha_1) - \tan(\alpha_2')} \tag{4}$$

$$l_{M1} = \sqrt{x^2 + y^2} \tag{5}$$

The mirror solution issue for α_1 can be solved by examining Δl for α_2 from equation (2). If this value is below zero, α_1 is correct. Otherwise, we have to negate α_1 (the mirror solution is true)[5]. The same applies for α_2 by examining Δl for α_1.

3 Overall Architecture and Processing Software

Figure 7 shows the overall architecture. To achieve the desired low-cost solution, a simple ATMega128 microcontroller[At11] has been used. The signal of the three microphones is sampled and converted by the integrated 10 Bit AD-converters of the microcontroller. The sampled data is collected in a buffer in real-time. The processing to retrieve the direction and distance of the sound source is then performed on this buffer. This concept detaches the time-sensitive data sampling from the computing-sensitive processing part.

[5] This applies only if the sound source is not located within the triangle given by the three microphone positions. However, in our experimental setup the three microphones form a single L-shaped rack of 0.245 meters side length. So the sound source cannot be located there.

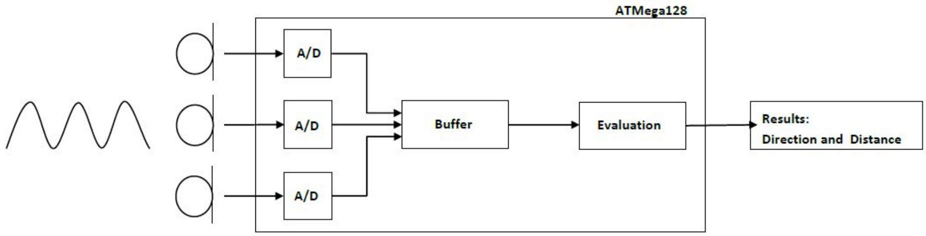

Fig. 7: Overall architecture

A classical approach to determine the phase shift between signals is the cross-correlation function [RG75]. For discrete real-valued functions it calculates to:

$$R_{xy}(\tau) = \sum_{\forall i} x(i) \cdot y(i + \tau)$$

The cross-corellation function calculates the similarity of two functions by different time shifts τ. If two function only differ by a time shift the cross-correlation function delivers a maximum value at this time-shift. However, the calculation of the cross-correlation function is complex. The computational complexity is $O(n^2)$, where n is the number of sampled values. To keep the calculation simple, a different approach has been used. The basic idea is to detect the zero-crossings of two signals and thereby determining their time shifts. This approach only produces the complexity $O(n)$, which is more suitable for the small microcontroller used. Furthermore, by using interpolation for the zero-crossing, the accuracy can be improved. The processing software consists of the following parts:

- *Minimum/maximum calculation of the sampled signal data*
  This is used to detect failures like signal clipping or a too low signal level.
- *Zero-level detection*
  To determine the zero-crossing of the sampled microphone signals, the zero-level for each microphone has to be calculated. This is done by averaging the sampled data values for each microphone[6].
- *Phase shift calculation*
  Using the detected zero levels, the phase shifts τ_1 and τ_2 between the microphones M_0, M_1 and M_1, M_2 (see Figure 5) are calculated. To increase accuracy, linear interpolation was used between the two values directly below and above the zero-level.
- *Calculation of the distance differences*
  Based on the phase shifts τ_1 and τ_2, the corresponding distance differences Δl_1 and Δl_2 are calculated using equation (1).

[6] An alternative approach would be to use the maximum/minimum values to calculate the zero-level. However, in our experiments the averaging approach has performed better.

- *Angle calculation*
  The angles α_1 and α_2 are calculated according to equation (2) and corrected as described to respect the possible mirror solution.
- *Distance calculation*
  The distance from microphone M_1 is calculated using equation (5).

4 Evaluation

To evaluate our approach, we have conducted several experiments. In this paper, we present the most important results. For more detailed information, see [Br16]. Firstly, the angle detection has been evaluated. The sound source has been placed at various angles and distances from the L-shaped microphone array. The upper part of Table 1 shows the result for a target angle of $\alpha_1 = 0°$. This means the sound source has been located in front of the microphone at an angle of 0° for α_1 at four different distances (40 cm, 55 cm, 100 cm, 150 cm). Of course, the corresponding target angle α_2 is distance-dependent and therefore slightly different for each of these distances (110°, 108°, 104°, 99°). Each measurement has been repeated three times and compared with the target values. It can be seen that the maximum inaccuracy for α_1 is 25°, while the maximum inaccuracy for α_2 is only 4°. The middle part of Table 1 shows the results for $\alpha_1 = 135°$ at the distances 20 cm, 50 cm, 100 cm and 150 cm. Here, the inaccuracy for both angles α_1 and α_2 is 14° at maximum.

All conducted experiments have shown that the measurement precision is best for angles close to 90° and decreases towards 0° and 180°. This behavior can also be explained mathematically. Looking at equation (2), the gradient of arctan is highest for angles towards 0° and 180° and therefore most sensitive to measurement imprecisions. Close to 90°, the gradient is lowest and therefore most tolerant to such imprecisons. Using equations (2), the angle inaccuracy α_{inacc} in relation to α, the measurement imprecicison for Δl (Δl_{innac}) and the microphone distance d can be calculated to:

$$\alpha_{inacc} = \left| \alpha - \arccos\left(\frac{d \cdot \cos(\alpha) + \Delta l_{innac}}{d} \right) \right|$$

Figure 6 shows the behavior of the angle inaccuracy for the microphone distance of our setup (0.245 meters) and an assumed measurement imprecison Δl_{innac} of 5 cm. It confirms the highest imprecisions for 0° (and 180° respectively) and a very low imprecision for 90° (= 1.57rad). The L-shape of our microphone array guarantees that at least one of the two angles α_1 and α_2 is in the range of low imprecision. Therefore, our approach performs very good at direction detection.

For distance detection, the results are different. Here, both angles are necessary and mostly one angle is the high imprecision range. Therefore, distance measurement is more sensitive to measurement inaccuracies. The lower part part of Table 1 confirms this assumption. The sound source has been placed in an angle of $\alpha_1 = 90°$ in four different distances (32 cm, 52 cm, 100 cm and 150 cm). Each measurement has been repeated 3 times and the computed values have been compared to the target distances. The table shows that, even though most of the values are quite precise, the possible measurement error can also be large. Exemplarily, the third measurement for the distance of 100 cm is nearly four times the real value. This is caused by an angle imprecision of only 10° in α_2. It can be seen that already small angle imprecisions can lead to large distance imprecisions. To sum it up, while the direction detection works very well the distance measurement is more an estimation than a precise result.

Target α_1	Measured α_1			Target α_2	Measured α_2			Distance
0	0	25	20	110	113	111	114	40
0	15	14	11	108	110	111	112	55
0	0	0	0	104	102	102	101	100
0	8	0	0	99	101	100	99	150
135	123	127	127	-135	-149	-148	-146	20
135	126	130	128	-135	-144	-145	-145	50
135	131	132	133	-135	-138	-139	-141	100
135	130	130	128	-135	-138	-140	-137	150

Target distance	Measured distance			Target distance	Measured distance		
32	43	51	48	52	49	45	50
100	77	79	375	150	92	99	137

Tab. 1: Evaluation results for angle and distance measurement (angles in degree, distances in cm)

5 Related Work

There exist several approaches dealing with the problem of sound source localization. These approaches are either hybrid solutions (combining radio and acoustic signal runtime [MB08]), based on an amplitude analysis of the acoustic signal or on phase shift detection like the approach presented in this paper. Since hybrid solutions are out of focus and the approaches based on amplitude analysis mostly deliver poor results [To76], we will exemplarily present two approaches using phase shift detection. [To76] focuses on the locatization of defects in materials by an array of three acoustic sensors. The basic idea is that defects in materials cause runtime differences of acoustic waves. The localization is done similar to the exact mathematical approach presented in the beginning of Section 2. Therefore this solution is very complex and suffers from ambiguities. This is the reason why we rely on an aproximation. Furthermore, the sound source should be placed in the triangle given by the acoustic sensors (inside measurement), while in our approach the sound source is located outside this triangle (outside measurement). Additionally, [To76] only deals with the geometric problem while techniques to detect the underlying phase shifts are not in the focus of this paper. [SMM97] presents an acoustical localization in three-dimensional space using crosspower phase-spectrum-analysis. This approach mainly focuses on detecting phase shifts. The geometric problem is of low relevance in this paper.

A main goal is the localization of a speaker in a room. Similar to [To76], but different to our approach it is supposed that the sound source is located within the space spanned by the microphones. 8 microphones have been used for the experiments. Furthermore, the spectral and correlation analysis require much more computational power than our approach.

6 Conclusions

This approach has investigated the possibilities to locate an acoustic sound source using low-cost hardware. The mathematical analysis and evaluation results have shown that the direction of the sound source can be detected with good precision. The distance measurement is more imprecise and is rather an estimation than a precise measurement. However, for the envisioned field of application like detecting the siren of a police car or a speaking person by an assistance robot, the direction detection is more important than the distance measurement. An assistance robot for example needs precise direction information for steering towards a person while for distance measurement an estimation is usually sufficient. Stopping in front of the person is mainly handled by obstacle detection sensors like supersonic range finders, laser scanners or cameras. However, it is possible to improve the distance measurement. One possiblity is to increase the measurement accuracy for the phase shift by increasing the sampling frequency. This would require the use of external analog/digital converters, since we have already used the maximum sample frequencey of the internal converters of the microcontroller for our experiments. Another possibility is to use more microphones. Using e.g. a T-shaped array of four microphones or two L-shaped arrays of three microphones each would allow to get two angles in high precision range and therefore a more precise distance calculation. This will be the topic of future work.

References

[At11] Atmel: ATMega128 and ATMega128L Datasheet, Atmel Cooperation, http://www.atmel.com/images/doc2467.pdf. June 2011.

[Br16] Brinkschulte, Melanie: Ortung einer sinusförmigen Schallquelle mithilfe eines Mikrocontrollers, Bachelor Arbeit, Fachbereich Informatik und Mathematik, Goethe Universität Frankfurt am Main. April 2016.

[MB08] Mühlberger, C.; Baunach, M.: Tab WoNS: Calibration Approach for WSN based Ultrasound Localization Systems, 7. GI/ITG KuVS FG Drahtlose Sensornetze. September 2008.

[RG75] Rabiner, Lawrence R.; Gold, Bernard: Theory and Application of Digital Signal Processing, Prentice-Hall. 1975.

[SMM97] Svaizer, P.; M.Matassoni; M.Omologo: Acoustic Source Location in a Three-Dimensional Space Using Crosspower Spectrum Phase, ICASSP-97 IEEE International Conference on Acoustics, Speech, and Signal Processing. 1:231–234, April 1997.

[To76] Tobias, A.: Acoustic-Emission Source Location in Two Dimensions by an Array of Three Sensors, Non-Destructive Testing, Elsevier Publishers. Februar 1976.