

Sound Source Localization in Wide-range Outdoor Environment Using Distributed Sensor Network

Mohammad Mahdi Faraji, Saeed Bagheri Shouraki, Ensieh Iranmehr, and Bernabe Linares-Barranco, *Fellow, IEEE*

Abstract—Sound source localization has always been one of the most challenging subjects in different fields of engineering, one of the most important of which being tracking of flying objects. This paper focuses on sound source localization using fuzzy fusion and a beamforming method. It proposes a new fuzzy-based algorithm for localizing a sound source using distributed sensor nodes. Eight low-cost sensor nodes have been constructed in this study each of which consists of a microphone array to capture sound waves. Each node is able to record audio signals synchronously on an SD card to evaluate different algorithms offline. However, the sensor nodes are designed to be able to estimate the location of the sound source in real-time. In the proposed algorithm, every node estimates the direction of the sound source. Moreover, a calibration algorithm is used for extracting the orientation of sensor nodes to calibrate the estimated directions. The calibrated directions are fuzzified and then used for localizing the sound source by fuzzy fusion. An experiment was designed based on localizing a flying quadcopter as a moving sound source to evaluate the performance of the proposed algorithm. The flying trajectory was then estimated and compared with the target trajectory extracted from the GPS module mounted on the quadcopter. Comparing the estimated sound source with the target location, a mean distance error of 6.03 m was achieved in a wide-range outdoor environment with the size of $240 \times 160 \times 80\text{ m}^3$. The achieved mean distance error is reasonable regarding the mean precision of the GPS module. The practical results illustrate the effectiveness of the proposed approach in localizing a sound source in a wide-range outdoor environment.

Index Terms - Sound source localization. Data fusion. Fuzzy logic. Distributed sensor network. Beamforming method.

I. INTRODUCTION

Due to the mechanical nature of sound waves, sound source localization (SSL) can be used in many areas, including localizing targets indoors, outdoors and even underwater. Therefore, sound source localization has many applications in several fields of engineering, such as biological sound studies [1], smart conference rooms [2], [3], traffic monitoring [4], underwater acoustical source localization [5], [6], [7], [8], speaker localization [9], and military industries [10] such as tracking of flying objects [11], sniper fire localization, and gunshot localization.

In order to find the direction of a sound source, two basic methods are used. The first method is based on time difference of arrival (TDOA) [12], [13], [14] and the second one is the beamforming method [15], [3], [16]. In this paper, a beamforming algorithm is used to estimate the direction of

Mohammad Mahdi Faraji, Saeed Bagheri Shouraki and Ensieh Iranmehr are with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: mmfaraji@ee.sharif.edu, bagheri-s@sharif.ir, eiranmehr@ee.sharif.edu)

Bernabe Linares-Barranco is with the CSIC Instituto de Microelectrónica Sevilla (IMSE-CNM), 41012 Seville, Spain (e-mail: bernabe@imse-cnm.csic.es).

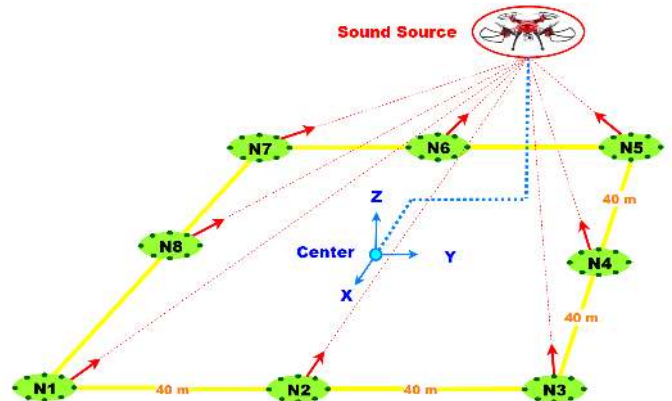


Fig. 1: 3D wide-range test environment for sound source localization. A quadcopter is considered as a sound source. There are eight distributed sensor nodes in the test environment, located on the perimeter of a square with the side of 80 m . The directions of the sound source viewed from the location of each sensor node are depicted using red arrows.

sound source. In addition to works which focus on finding the direction of a sound source [17], [18], [13], there are several works which focus on estimating its coordinates [12], [14]. Despite the fact that most of these works study indoor sound source localization [18] for different applications [19], [2], [3], [9], there are a few works which focus on outdoor sound source localization [11], [7], [8]. Hence, this paper focuses on finding the coordinates of a sound source in a wide-range outdoor environment. On the other hand, the use of quadcopter has recently grown exponentially, and nowadays it is possible to easily get a quadcopter with the appropriate flight capabilities at a low cost. Despite the facilities that they provide, these flying objects can cause many problems. For instance, they can cause disturbance in airport performance and even be used for destructive operations. Therefore, there is a growing need in localizing and detecting these flying objects. Hence, in this paper, a quadcopter is considered as a sound source. Quadcopters and other flying objects emit sound waves while flying because of their propellers. In order to capture the sound waves released from a quadcopter, several designed sensor nodes are distributed in the environment. Each sensor node designed in [13] is equipped with a MEMS microphone array and is used to find the direction of the sound source (quadcopter). Different literatures show that using microphone arrays in localizing sound sources is rather common [20], [21], [19], [16], [2], [22], [23], [24], [17]. In this paper, each node is able to sample the sound waves using an array of microphones. The audio signal is then used to estimate the direction of the sound source in each sensor node. In this paper, in order to find the exact location of the sound source, the estimated directions are fused together using fuzzy fusion algorithm.



Fig. 2: The hardware of the designed sensor node. It is made up of a main processing board and a circular microphone array.

In this paper, like many other previous approaches such as [25], eight sensor nodes distributed in the environment are used to localize a sound source in a wide-range environment. Fig. 1 illustrates an overview of the proposed system. As can be seen in this figure, there are eight distributed sensor nodes that are located on the perimeter of a $80 \times 80 m^2$ square. Each sensor node is equipped with a microphone array and a processing unit which is able to process the receiving sound waves and extract the direction of the sound source in real-time. After finding the direction of the sound source in each sensor node, these directions are fused together to find the actual location of the sound source. This paper mainly focuses on estimating the coordinates of the sound source using distributed sensor nodes.

In this paper, the fusion of the estimated directions is accomplished by means of fuzzy reasoning method. Fuzzy set theory, introduced in 1965 by L.A. Zadeh [26], has progressed to become a powerful mathematical theory over the last few decades. This paper combines fuzzy fusion with signal processing methods. Fuzzy fusion has also been used successfully in many applications such as multirobot object localization problem [27]. In this paper, we have used fuzzy logic to fuse the estimated directions of the sound source together which results in an accurate estimation of the sound source location.

The rest of this paper is organized as follows: The proposed sensor architecture is introduced in Section II. It is built to create an audio database for research applications as well as real-time localization of the sound source. Section III represents the beamforming algorithm for estimating the sound source direction. Section IV represents the proposed fuzzy localization algorithm for estimating the coordinates of the sound source using the directions estimated by sensor nodes. In Section V, the proposed method is evaluated considering a test scenario in outdoor environment using eight sensor nodes. Section VI discusses how to increase the precision in future works. Finally, we conclude the paper in Section VII.

II. HARDWARE STRUCTURE OF A SENSOR NODE

This section introduces the designed sensor's hardware. The sensor node is battery powered and is designed as small as

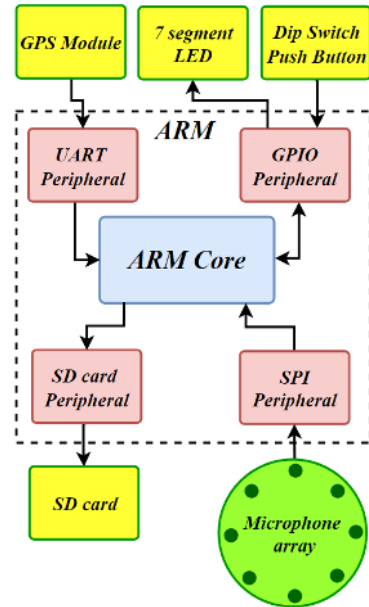


Fig. 3: The proposed sensor node architecture. The designed microphone array, an ARM micro controller, an SD card and a GPS module are used in the designed sensor nodes. The SD card records synchronous global time-stamped audio data which is provided via the GPS module.

possible so that it would be useful for portable applications such as sound source localization in real-time. It can also be used to create an audio database for research purposes which can be used to evaluate and compare different sound source localization algorithms.

The designed sensor node has eight digital MEMS microphones placed on a circular PCB with a diameter of $16 cm$ which is placed at a specific distance from the main $9.5 cm \times 9.5 cm$ square board to prevent sound reflections, as shown in Fig. 2. To execute sound source localization algorithms in real-time, an ARM micro controller is used in the designed sensor node. The sensor node can be used as a standalone sensor or as a part of a distributed sensing application setup.

Just like many other array processing applications, in the designed sensor node it is essential for all the microphones to sample sound signals synchronously. Hence, ICS-52000 MEMS microphones, which are high-quality digital microphones with a time division multiplexing (TDM) audio interface [28], are used in the designed sensor nodes. An array of up to 16 microphones can be connected directly to digital processors with no need for the system to have an audio codec because of the TDM audio interface.

The architecture of the designed sensor node is shown in Fig. 3. It consists of a microphone array for sampling the sound present in the environment, a global positioning system (GPS) module for synchronization and positioning, an SD card for recording the sound, an ARM micro controller as a processor, and several general-purpose input/outputs (GPIOs) such as seven segment displays, light-emitting diodes (LEDs), five push buttons, and a dip switch.

The audio signals of the microphone array are received by the ARM micro controller via serial peripheral interface (SPI)

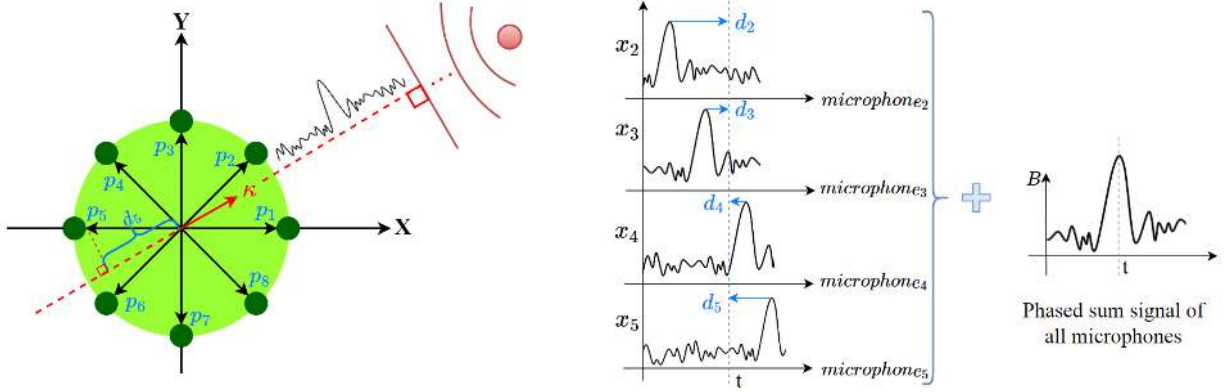


Fig. 4: A scheme of delay and sum beamforming method. In a specific direction of κ , the signal received from each microphone in the array is shifted according to its delay time. Then, they are all added together to focus on that direction. The unitary direction vector of a far-field sound source is defined by κ . d_m shows the delay time of the m^{th} microphone. Signals coming in from the focused direction are amplified once all the shifted signals are added. The beamformed signal for direction κ is then computed which is shown by B .

and are written on an SD card. It is necessary for the audio samples to be stamped by synchronous global time. Hence, an L86 GPS module is used [29] in the sensor node to position the sensor node. It is connected to the micro controller via universal asynchronous receiver-transmitter (UART). Using a pulse per second (PPS) signal received from the GPS module, audio samples are stamped by synchronous global time. A Li-ion battery with a capacity of 2600mA/hr is used in the sensor node to provide the required energy. By using this chargeable battery, the sensor node can run with full performance up to 26 hours on a single charge which is an acceptable duration for research purposes. To facilitate debugging, the status of the sensor node is displayed by LEDs while track number, error numbers and node number are displayed on 7-segment displays. Also, to provide a user input interface, push buttons and a dip switch are used in the sensor node.

III. BEAMFORMING ALGORITHM TO ESTIMATE SOUND SOURCE DIRECTION IN EACH SENSOR NODE

In this paper, eight sensor nodes are used for sound source localization. Each sensor node is responsible for estimating the direction of the sound source and then by fusing the estimated directions, the position of the sound source is achieved. This section presents a beamforming method for estimating sound source direction in each sensor node. To do so, the probability of sound source presence for each hypothetical direction has to be computed. The sound source can then be considered to be located at the direction with maximum probability. An angular region of interest (ROI) is defined and is then quantized to a certain specific number of points. The number of defined points depends on the amount of processing power available. Each angular ROI point associated with a specific direction, depends on the angle precision. This section presents the beamforming algorithm used for estimating sound source direction by computing the probability of sound source presence in the quantized angular ROI.

Delay and sum beamforming (DSB) method which is used in this paper to find the direction of a sound source, is the simplest existing beamforming method [30]. By using DSB,

signals coming from a specific direction are amplified while signals coming from all other directions are suppressed. A scheme of delay and sum beamforming method is shown in Fig. 4. The beamformed signal is computed for each hypothetical sound source direction by delaying every microphone signal and then adding them all together. The time delay of each microphone signal can be easily computed considering its position in the array and the hypothetical sound source direction which is shown by κ in Fig. 4. For simplicity, this figure shows sound source localization in 2D environment. However, this direction (κ) is defined as 3D in this paper.

In Fig. 4, a far-field sound source is considered which may make array analysis easier. In this paper, the distance between the far-field sound source and the microphone array is ten times greater than the size of the microphone array. Hence, the distance between the sound source and the microphone array is not entered in the computation of the sound source direction. In this figure, κ shows the unitary direction vector, and d_m shows the delay time of the m^{th} microphone with respect to the center of the microphone array. If the position vector of the m^{th} microphone is shown by p_m , $d_m(\kappa)$ is computed by projecting p_m onto unitary vector κ using Eq. 1, in which c is the sound speed. Both the position of the m^{th} microphone in the array and the desired focus direction affect d_m .

$$d_m(\kappa) = \frac{p_m \cdot \kappa}{c} \quad (1)$$

Fig. 4 also shows the audio signals of the four microphones. It should be noted that their times of arrival are different due to their positions in the array. As can be seen in this figure, there is a peak in the audio signal. Let us suppose that the peak reaches the center of the microphone array at time t . Hence, some microphones such as 1, 2, 3 and 8 receive it some samples earlier than it arrives at the array center while the others such as 4, 5, 6 and 7 will receive it some samples later. The time at which the signal arrives at the center of the array is shown in the figure by a blue dash-line.

The signal energy for each angular ROI point in DSB [31] can be computed by delaying the audio signals of the microphones and then adding them all together. Each angular ROI point indicates a specific unitary direction vector κ . To

Algorithm 1: Estimating the Direction of a Sound Source

Initialization: define an angular ROI with N points (each angular ROI point declares a specific direction);

Output: 8 directions estimated by 8 sensor nodes

```

for (every sensor node)  $S = 1$  to 8 do
  for (every angular ROI point)  $P = 1$  to  $N$  do
    Compute the beamformed signal energy
    ( $E(B(\kappa_P))$ ) using Eq. 2;
  end
  Find the direction with the maximum beamformed
  signal energy ( $\kappa_{max}^S$ ) using Eq. 3;
end

```

compute the signal energy for each direction, Eq. 2 is used, where x_m is the m^{th} microphone's signal and d_m is the amount of delay applicable to the m^{th} microphone which is computed using Eq. 1. The signal focused on direction κ is shown by $B(\kappa)$ in this equation. The m^{th} microphone's audio signal is shifted by the value of d_m . The shifted signals of all microphones are added together to focus on direction κ .

$$B(\kappa) = \sum_m x_m(t - d_m(\kappa)) \quad (2)$$

$$\kappa_{max} = \underset{\kappa_P}{\operatorname{argmax}} E(B(\kappa_P)) \quad (3)$$

As mentioned above, each angular ROI point declares one specific direction. After computing the signal energy of all angular ROI points, a direction with the maximum beamformed energy (E) is obtained using Eq. 3 and is considered as an estimated sound source direction. Algorithm 1 represents the steps of finding the direction of a sound source from the view of each sensor node.

IV. FUZZY FUSION OF THE ESTIMATED DIRECTIONS

In this paper, for sound source localization, the fuzzy belief of sound source presence is estimated for each hypothetical location. The sound source location can then be considered at the point at which the estimated belief value is maximum. Due to the limitations in processing resources, the search area or region of interest (ROI) is usually limited to a specific number of points. A ROI can vary depending on the application. In the case of tracking flying objects [11], it is desirable to localize the sound source in 3D space while in the case of traffic monitoring [4], it is sufficient to localize it in 2D space. Depending on the amount of processing power available, the ROI has to be quantized to a specific number of points.

This section explains the proposed sound source localization algorithm which is based on fuzzy fusion of estimated directions. Algorithm 2 represents the proposed method for estimating the location of the sound source. Two types of ROIs were defined in this study: one is angular ROI for direction estimation which was defined in Section III and the other is spatial ROI for localization which is defined in this

section. The proposed algorithm is able to estimate the belief of sound source presence in the quantized spatial ROI points. In order to determine the location of the sound source, the fuzzy beliefs of the directions estimated by the sensor nodes for each hypothetical spatial ROI point are extracted. Then, these fuzzy beliefs are fused and eventually the location of the sound source is estimated using a defuzzifier. This section consists of three subsections. Subsection IV-A provides the explanations on how to fuzzify the direction estimated by each sensor node. For each spatial ROI point, fuzzy beliefs of the estimated directions are extracted. The method of fusing fuzzy beliefs is explained in Subsection IV-B and finally the defuzzification method used to estimate the location of the sound source is explained in Subsection IV-C.

A. Fuzzifying the estimated directions

The beamforming algorithm mentioned in Section III, computes the amount of focused sound signal energy for each direction in each time window. These focused sound signal energies can be different per time window for each sensor node; therefore, in this paper, the focused sound signals are decided to be described as fuzzy beliefs. The following provides explanations on how to assign a belief degree to a focused sound signal. According to the beamforming algorithm, a direction in which the amount of focused audio signal energy is greater is considered as a sound source direction in each time window. It can be concluded that the amount of signal energy at each direction can represent the belief of sound source presence at that direction. Since the amount of signal energy depends on the intensity of the received sound signal, the signal energy needs to be normalized. If the maximum and minimum amount of energy for each sensor node are presented by E^{max} and E^{min} respectively and κ_{max} represents the direction with the maximum focused energy, $\mu(\kappa_{max})$ is computed using Eq. 4 which represents the belief of sound source presence at direction κ_{max} . The direction with the maximum energy in the time window t , which is represented by κ_{max}^t , is computed using Eq. 5. It should be noted that P in κ_P declares the angular ROI point number. $E(B(\kappa_P^t))$ represents the focused signal energy at the direction κ_P and in the time window t .

According to Eq. 4, if E^{max} and E^{min} are close to each other in value, the maximum value of fuzzy belief is decreased. This situation indicates the uncertainty of the estimated direction in the sensor node. On the other hand, if the amount of E^{max} is much greater than the amount of E^{min} , it means that the direction is estimated with great confidence.

The sound source direction estimated by each sensor node is a crisp direction. A triangular membership function is assumed around the crisp direction to fuzzify it. A fuzzy set Λ on κ_P is defined by the triangular membership function $\mu_\Lambda: \kappa_P \rightarrow [0, \mu(\kappa_{max})]$. Fig. 5 shows the triangular membership function around κ_{max} . A fuzzy membership function around the estimated crisp direction for a time window t can be considered using Eq. 6 in which $2 \times w$ is the base size of the triangular membership function. By using this equation, the belief of sound source presence at every direction is computed.

Algorithm 2: Estimating the Location of a Sound Source

Initialization: define a spatial ROI with M points
for (every spatial ROI point) $P = 1$ to N **do**
 for (every sensor node) $S = 1$ to 8 **do**
 Compute the direction of assumed point P from
 the view of the S^{th} sensor node (D_P^S) using
 Eq. 7;
 end
end
Input: 8 directions estimated by 8 sensor nodes
Output: location of the sound source

for every time window **do**
 for (every sensor node) $S = 1$ to 8 **do**
 Fuzzify the S^{th} estimated direction using Eq. 6
 which leads to obtaining the S^{th} fuzzy
 membership function (Λ^S);
 end
 for (every spatial ROI point) $P = 1$ to N **do**
 for (every sensor node) $S = 1$ to 8 **do**
 Find the membership value of the direction
 D_P^S from the membership function Λ^S
 and obtain $\mu_{\Lambda_S}(D_P^S)$;
 end
 Fuzzy fusion of the membership values which
 leads to obtaining the truth value of point P
 using Eq. 12 ($\hat{\mu}^P$);
 Find the point of ROI with the maximum truth
 value and consider it as the sound source
 location using Eq. 14;
 end
end

The fuzzy belief of sound source presence at direction κ_P and in the time window t is represented by $\mu_{\Lambda}(\kappa_P^t)$. In other words, $\mu_{\Lambda}(\kappa_P^t)$ is the membership value of the direction κ_P^t in the fuzzy set.

$$\mu(\kappa_{max}) = \frac{(E^{max} - E^{min})}{E^{max}} \quad (4)$$

$$\kappa_{max}^t = \underset{\kappa_P^t}{\operatorname{argmax}} E(B(\kappa_P^t)) \quad (5)$$

$$\mu_{\Lambda}(\kappa_P^t) = \begin{cases} \frac{|\kappa_P^t \cdot \kappa_{max}^t - 1 + w|}{w} \times \mu(\kappa_{max}) & |1 - \kappa_P^t \cdot \kappa_{max}^t| < w \\ 0 & |1 - \kappa_P^t \cdot \kappa_{max}^t| > w \end{cases} \quad (6)$$

As mentioned above, directions estimated by sensor nodes are fuzzified based on the proposed fuzzifier which is shown in Fig. 5. Each fuzzified direction is presented by two parameters: the direction of the sound source estimated in each sensor node (the direction with maximum beamformed energy) and its fuzzy belief (computed using Eq. 4). Hence, it only needs to send these two parameters at any time. This leads to the transmission of the minimum amount of data between sensor nodes. As mentioned above, the direction of the sound source

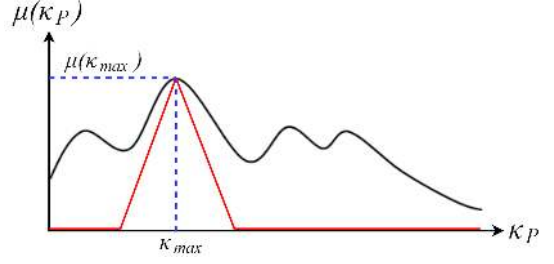


Fig. 5: The proposed fuzzifier. The black curve shows the beamformed signal energy. The proposed fuzzifier is a triangular membership function (μ_{Λ}) around the direction with the maximum beamformed energy (κ_{max}) which is shown by a red curve.

estimated in each sensor node is a crisp direction and is fuzzified by a triangular membership function in order to convert it to a fuzzy direction. Using this triangular membership function may result in robustness in direction estimation. Direction estimation error is very common due to the presence of noise and angular ROI quantization. Therefore, the proposed fuzzifier compensates for the direction estimation error. It can be concluded that these equations can present a proper fuzzy description from the result of beamforming algorithm.

In this study, eight sensor nodes are used for sound source localization. Therefore, eight fuzzy directions each associated with a sensor node are obtained in each time window. By fusing these fuzzy directions, the location of the sound source is estimated. The next subsection provides explanations on how these fuzzy directions are fused.

B. Fusion of the fuzzy beliefs of the estimated directions in 3D space

After obtaining the fuzzy directions according to beamforming algorithm for each time window, these directions have to be fused to estimate the position of the sound source in 3D space. For this purpose, a spatial ROI has to be defined and quantized into a specific number of points. The amount of fuzzy belief of sound source presence for each ROI point is then computed.

In order to estimate the location of the sound source in space, each spatial ROI point P has to be assumed as a sound source location and the directions of the assumed sound source with respect to every sensor nodes have to be computed once at the initialization step of Algorithm 2. The direction of an assumed point P from the view of the S^{th} sensor node is easily computed using Eq. 7. In this equation, D_P^S denotes the unitary direction vector which is computed based on the location of the S^{th} sensor node (L^S) and the location of assumed point P (L^P). The belief of the S^{th} fuzzy direction (Λ_S) for a hypothetical ROI point P is represented by $\mu_{\Lambda_S}(D_P^S)$. To estimate the location of the sound source, the beliefs of eight fuzzy directions for every ROI point have to be obtained first. Then for each ROI point, these eight beliefs have to be fused by fuzzy triangular conorms (T-conorms or S-norms) using Eq. 8 in which $\hat{\mu}^P$ represents the belief of sound source presence at point P .

$$D_P^S = \frac{\overline{L^S L^P}}{|L^S L^P|} \quad (7)$$

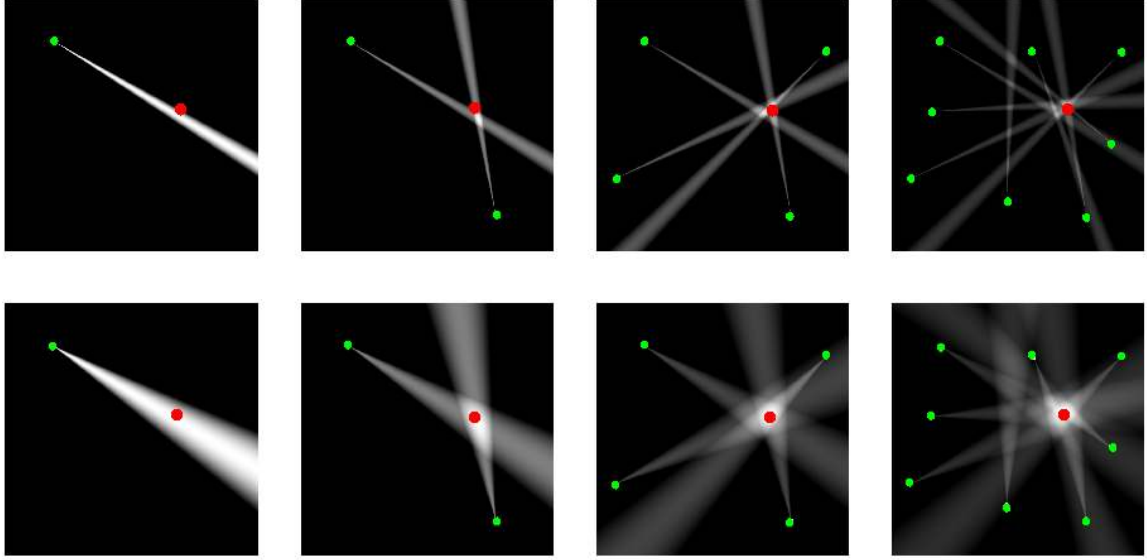


Fig. 6: Fusion of the fuzzy beliefs considering different numbers of sensor nodes and using the proposed fuzzifier with two different base sizes of triangular membership function. The position of the sensor nodes and the sound source are obtained by the GPS module. Green circles show the position of the sensor nodes and the sound source position is shown by a red circle. First row: fusion results when using the proposed fuzzifier with a very small base size (very close to crisp) considering 1, 2, 4 and 8 sensor nodes. Second row: fusion results obtained from using the proposed fuzzifier with a larger base size considering 1, 2, 4 and 8 sensor nodes. Each node estimates the direction of the sound source. The estimated direction is then fuzzified which is shown by a bright beam. All the directions estimated by sensor nodes are fused together to determine the position of the sound source. The brighter area represents higher belief value of sound source presence at that position. Using more sensor nodes will lead to obtaining more accurate results. If some of the sensor nodes, as shown in the last column, estimate the direction of the sound source with considerable error, the others compensate for the effect of the wrong estimated direction.

$$\hat{\mu}^P = \prod_{S=1}^8 (\mu_{\Lambda_S}(D_P^S)) \quad (8)$$

Different types of t-conorms [32] are frequently used in different applications. Let us suppose that for an element x of X , the membership degree in the fuzzy set A and B is expressed by $\mu_A(x)$ and $\mu_B(x)$ respectively. One of the t-conorms is maximum or the Godel t-conorm as shown in Eq. 9. Another t-conorm is bounded sum or the Lukasiewicz t-conorm, shown in Eq. 10. The third commonly used t-conorm is product t-conorm or probabilistic sum, shown in Eq. 11. Due to the independent effect of each node on localizing the sound source, sum based t-conorm is required in this study. Moreover, to avoid saturation in the output membership function, product t-conorm is used for fuzzy fusion of direction beliefs for every ROI point. Eq. 12 represents the computation process of sound source presence belief at the P^{th} ROI point.

$$\mu_{A \cup B}(x) = \max(\mu_A, \mu_B) \quad (9)$$

$$\mu_{A \cup B}(x) = \min(\mu_A(x) + \mu_B(x), 1) \quad (10)$$

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x) \quad (11)$$

$$\hat{\mu}^P = \frac{1}{\binom{8}{2}} \sum_{S=1}^7 \sum_{j=S+1}^8 (\mu_{\Lambda_S}(D_P^S) + \mu_{\Lambda_j}(D_P^j) - \mu_{\Lambda_S}(D_P^S) \times \mu_{\Lambda_j}(D_P^j)) \quad (12)$$

In fact, in the proposed localization algorithm, the fuzzy belief of sound source presence at each ROI point $\hat{\mu}^P$ is obtained by fusing the fuzzy directions using the t-conorm shown in Eq. 12. Therefore, a fuzzy set on the spatial ROI points (P) with a membership function $\hat{\mu}: P \rightarrow [0, 1]$ is obtained where each spatial ROI point is mapped to a value between 0 and 1. As mentioned in the previous subsection, the estimated directions are presented by triangular fuzzy membership functions. To illustrate the fusion of the fuzzy directions, a simulation was designed. In this simulation, the sensor nodes were assumed to be located in a $100 \times 100 m^2$ area. The simulation results are shown in Fig. 6. The figure shows the fusion results of the fuzzified directions. The fusion results, which lead to estimating the location of the sound source, were drawn for different numbers of sensor nodes. Moreover, two different base sizes of triangular membership functions were considered in this simulation: one with a very small base size, very similar to crisp directions, and one with a larger base size used to show the fusion of fuzzy directions in sound source localization. As can be seen in this figure, the fusion of fuzzy directions estimates the location of the sound source better than the fusion of crisp directions. Brighter areas indicate higher belief of sound source presence in contrast to darker areas which indicate lower belief of sound source presence. In other words, the sound source presence in 3D environment is represented by fuzzy beliefs each of which is associated with a point of spatial ROI. However, the fuzzy output of sound source localization algorithm is not acceptable for many systems and therefore, it is necessary to determine an exact point of ROI as a sound source location. Hence, the

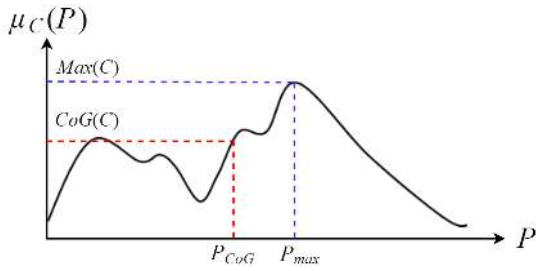


Fig. 7: Center of gravity defuzzification method in comparison with maxima defuzzification method. The maximum of fuzzy set C is represented by $Max(C)$ while the CoG of fuzzy set C is represented by $CoG(C)$. The ROI point with the maximum belief is shown by P_{max} .

fuzzy membership function has to be defuzzified in order to determine a point of ROI as a sound source location. In the next subsection, we explain the defuzzification method.

C. Defuzzification of fusion results

The purpose of defuzzification is to convert the fuzzy membership function to a crisp value in order to estimate the exact position of a sound source. There are several defuzzification methods [33], the most well-known of which is center of gravity (CoG) defuzzification method which computes the center of gravity of the area under the membership function. If the fusion result is represented by fuzzy set C , then CoG is computed using Eq. 13 in which $\hat{\mu}_C(p)$ indicates the membership degree of point P in ROI. Another well-known type of defuzzifiers is the maxima methods. The maxima methods have the common property of selecting an element of fuzzy set as defuzzication value. The maximum point of the fuzzy membership function is obtained using Eq. 14. As seen in this equation, a point in ROI with the maximum membership degree is considered as a sound source location. In other words, a defuzzifier based on maximum belief outputs a point while a CoG defuzzifier outputs a weighted average of the fuzzy beliefs of ROI points. Fig. 7 illustrates the way these two types of defuzzifiers work for a sample fuzzy set. Depending on the application, we can decide on the more suitable type of defuzzifier.

$$CoG(C) = \frac{\sum_P P \times \hat{\mu}_C(P)}{\sum_P \hat{\mu}_C(P)} \quad (13)$$

$$Max(C) = \underset{P}{argmax} \hat{\mu}_C(P) \quad (14)$$

In the application of sound source localization, the maxima-based defuzzifier is more appropriate than CoG defuzzifiers due to the acoustical noise existence in the environment. This noise, leading to the wrong directions computed by some sensor nodes, may have a negative effect on the fuzzy membership function as seen in the last column of Fig. 6. When one of the sensor node makes a mistake in estimating the direction of sound source, its center of gravity can be changed remarkably while using maxima-based defuzzifier, the effect of noise is eliminated. Therefore, using maxima-based defuzzifier leads to the more robust sound source localization than the CoG defuzzifier. In this paper, the maxima-based defuzzifier

is used in which the location of sound source is estimated as a point in spatial ROI where most of the sensor nodes confirm the presence of sound source on that point.

The maxima-based defuzzifier faces ambiguity when more than one maximum is occurred. Due to the fusion of eight fuzzy directions computed by eight sensor nodes to obtain the fuzzy belief of sound source presence ($\hat{\mu}$) according to Eq. 12, the maximum belief is rarely occurred at more than one spatial ROI point. Despite maxima-based defuzzifier faces ambiguity in this rare situation, we can simply overcome this challenge using the motion model of the sound source. Suppose more than one point of spatial ROI at time t has the maximum belief. In this case, the position of the sound source can be estimated considering the positions of the sound source at two previous times $t-1$, $t-2$. Assuming that the time-step for estimating the sound source position is the same for different times, the estimated position at time t (P_e^t) can be calculated using Eq. 15. In this equation, P^{t-1} and P^{t-2} denote the actual position of the sound source at time $t-1$ and $t-2$, respectively. Also, $(P^{t-1} - P^{t-2})$ represents an estimate of the amount of the source movement in a time-step. In other words, using Eq. 15, the estimated position at time t can be achieved by adding the position at $t-1$ to this amount of movement. After estimating P_e^t , a point of spatial ROI among the ones that have maximum belief with the least distance to P_e^t is chosen as the sound source location.

$$P_e^t = P^{t-1} + (P^{t-1} - P^{t-2}) \quad (15)$$

It should be noted that the computational cost of defuzzification based on maximum is less than that of CoG. According to the explanations provided in this section, in this paper, the triangular fuzzifier, the product t-conorm and the maximum based defuzzifier were used to localize the sound source.

V. EXPERIMENTAL RESULTS

In this section, an experiment carried out to evaluate the proposed sound source localization algorithm is explained. In this experiment, eight sensor nodes were used and located on the perimeter of a square with the side size of 80 m as shown in Fig. 1. The exact position of each sensor node was extracted by the GPS module used in each sensor node. As mentioned in Section II, the microphone's audio signal which is stamped by synchronous global time and the location captured by the GPS are recorded on an SD card. Recording the data on an SD card makes it possible to evaluate and compare different sound source localization algorithms in offline. In this test, a flying commercial quadcopter (SYMA X8HG) was localized, as shown in Fig. 8. The quadcopter was equipped with a GPS module and also an SD-card to record its travel path, in order to evaluate the accuracy of the proposed localization algorithm. The sound pressure level (SPL) at one meter distance from the quadcopter was about 70 dB . The background noise in the outdoor test environment was about 34 dB SPL . In this test, the quadcopter is considered as a far-field sound source since the distance between the sound source and the sensor node is much greater than the array size. As mentioned in previous sections, two main steps have to be considered for

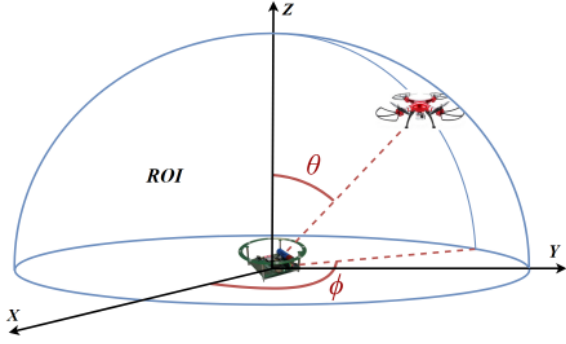


Fig. 8: Angular ROI definition for DE. Each sound source direction is equivalent to a point in the ROI. A quadcopter flies in a 3D outdoor environment and the designed sensor node captures its sound waves to estimate its direction. The direction of the quadcopter from the view of the sensor node is shown by polar angles θ and ϕ .

localizing the sound source: one is direction estimation (DE) for each sensor node and the other, is fusion of the estimated directions to find the sound source location (SSL).

For DE, the hemispherical surface shown in Fig. 8 was assumed as an appropriate angular ROI for the sensor node. Polar angles θ and ϕ in the ROI were also quantized into a specific number of points. Each point of angular ROI represents a direction of the sound source in space. As can be seen in Fig. 8, the ranges of polar angles θ and ϕ are limited to 90° and 360° , respectively. The number of ROI points depends on the angular resolution required. For example, if the resolution of polar angles is considered 5° , then the angular ROI would be quantized into $\frac{90}{5} \times \frac{360}{5} = 1296$ points.

After finding the directions estimated by sensor nodes, the directions have to be fused to estimate the location of the sound source. Therefore, an appropriate 3D spatial ROI was considered to estimate the position of the sound source in Cartesian coordinate system. Depending on the precision we need, the coordinates x , y and z were also quantized into a specific number of points. Each point of ROI space represents one hypothetical position for the sound source in the environment. In this practical experiment, the quadcopter was flying in a 3D space with the range of 240 meters at x coordinate, 160 meters at y coordinate and 80 meters at z coordinate. For example in 3D space quantization, if the space resolution is considered 10 m, the number of ROI points is $\frac{240}{10} \times \frac{160}{10} \times \frac{80}{10} = 3072$ for SSL. It is worth mentioning that the angular and spatial resolutions depend directly on the available sensor node processing resources.

To localize the sound source more accurately, we have tried to place each node horizontally towards the north. However, the angular deviance is unavoidable because of installation error. It is therefore necessary for every direction estimated by sensor nodes to be calibrated first. To calibrate the polar angles θ and ϕ estimated by each sensor node, a single tone sound source with high sound pressure level (SPL) was being played in 10 specific points of the environment and the sensor nodes captured this sound wave. Considering the point at which the single tone sound source is located and the position of the sensor nodes, the direction of the sound source for each sensor node is computed using Eq. 7. On the other hand, each sensor node estimated the direction of this sound

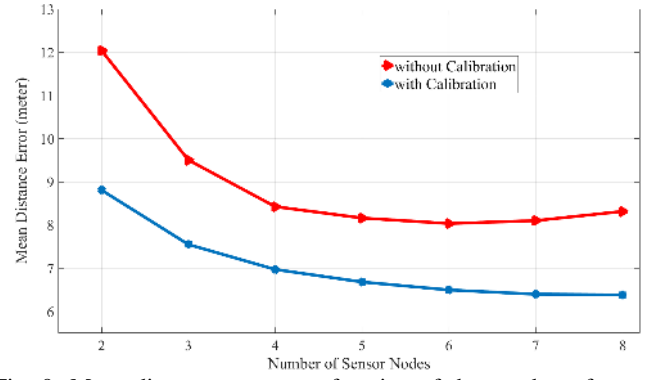


Fig. 9: Mean distance error as a function of the number of sensor nodes considering different calibration status. Using different number of nodes causes different mean distance errors. The mean distance error achieved when using a higher number of sensor nodes shows that the more the sensor nodes, the more accurate the results. The mean distance error achieved when using calibration is much less compared to when calibration is not used.

source using beamforming method. The difference between the computed direction and the direction estimated by each sensor node determined the angular deviation. By using the angular deviation, the calibration was performed. Then, the fuzzy fusion method which was explained in the previous section, was applied on the fuzzified calibrated directions and the location of sound source was estimated.

The mean distance between the estimated position and the position extracted by the GPS sensor mounted on the quadcopter is considered as an evaluation criterion. In the following, we show that the mean distance error is decreased when using calibration for localizing a sound source. In order to show the effect of calibration and also the number of sensor nodes on localization, a test was designed. In this test, a flying quadcopter was considered and the data were recorded on SD cards. Then, the mean distance error of localization was computed offline using different numbers of sensor nodes both with and without calibration. Fig. 9 shows the effect of calibration and the number of sensor nodes on sound source localization. The less the mean distance error, the more accurate the localization results. In this test, the spatial ROI resolution is considered 5 meters and the angular ROI resolution is considered 5° . As can be seen in this figure, the mean distance error by considering calibration is less than the mean distance error without considering calibration. Moreover, it can be seen that the mean distance error is generally decreased by increasing the number of sensor nodes. By considering calibration in this test, the best mean distance error achieved is 6.38 meters which is obtained when using 8 sensor nodes. However, the best mean distance error achieved without calibration is 8.04 meters which is obtained while using 6 sensor nodes. It means that without calibration using more sensor nodes may increase the mean distance error. While by using calibration, using more sensor nodes may reduce the mean distance error and increase the localization accuracy.

To evaluate the proposed sound source localization algorithm, we flew the quadcopter for about 10 minutes. By using the directions estimated by the sensor nodes, the location of

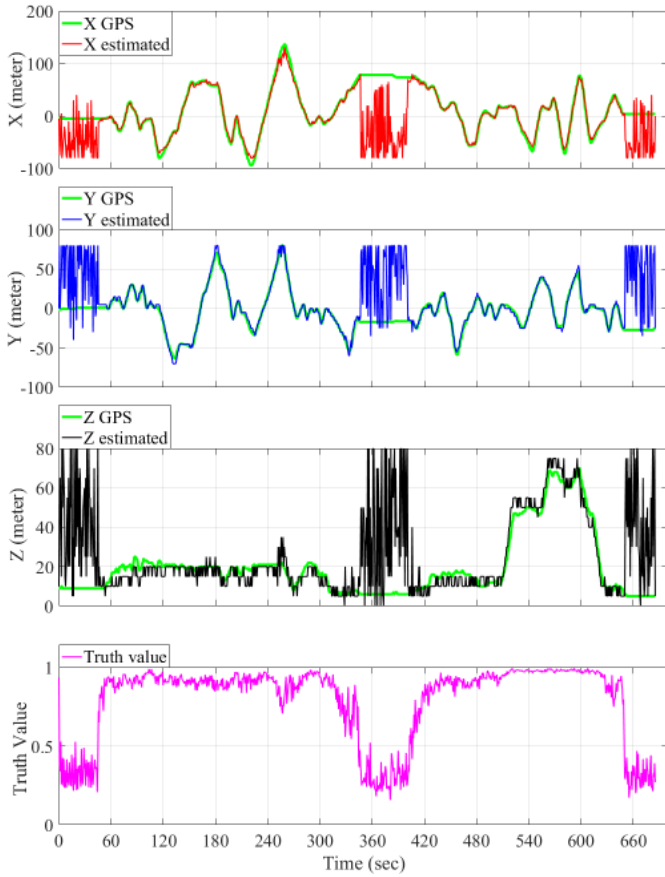


Fig. 10: Comparison between the estimated location (x, y, z) of the flying object and the GPS data for a time period of about 10 minutes. The first, second and third graph show the estimated x, y and z location of the flying quadcopter in comparison with the target location extracted from the GPS sensor respectively. The fourth graph shows the truth value of the estimated location which means that the belief of quadcopter presence at the estimated location. As can be seen, there are time pieces in the graphs which indicate a large difference between the estimated location and the location captured by the GPS module. These time pieces represent the times when the quadcopter is crashed and there is no sound emitted from the quadcopter. It can be seen that the truth value in these time pieces is less than 0.6 while it is more than 0.6 for other times. It means that the quadcopter location for the times with the truth value of more than 0.6 is estimated correctly or at least closely to the target location.

the quadcopter at each moment was estimated. The estimated location of the sound source can be at any point of spatial ROI. As mentioned before, the spatial ROI was quantized into the specific points depends on the spatial resolution. The location of the sound source in space is represented by Cartesian coordinates (x, y, z) . Both the estimated and the target location are drawn in Fig. 10 during the quadcopters's flights. The target location, which is called $(X \text{ GPS}, Y \text{ GPS}, \text{ and } Z \text{ GPS})$ and shown in green in this figure, is extracted by the GPS module mounted on the quadcopter. On the other hand, the proposed localization algorithm estimates the location of the quadcopter once in every second. The estimated locations (x, y, z) are shown in red, blue and black respectively. The figure also shows the truth value or belief of the estimated location which is shown in magenta. As can be seen in

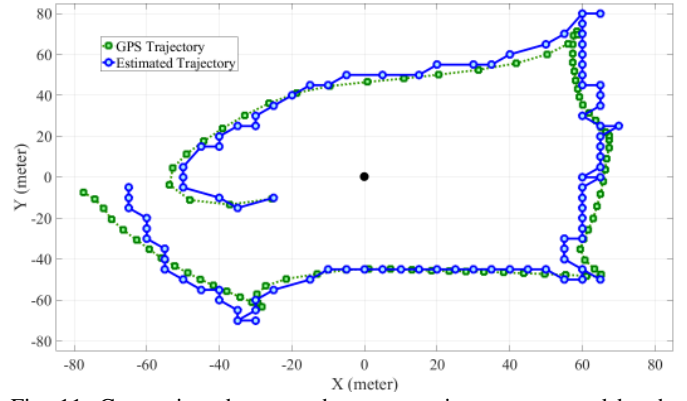


Fig. 11: Comparison between the target trajectory extracted by the GPS module and the estimated trajectory by the proposed algorithm. The estimated trajectory is shown in 2D space for simplicity while it is estimated in 3D space. When the quadcopter flies near the center of the sensor nodes shown by a circle, the estimated trajectory is very similar to the trajectory extracted by the GPS sensor. It is also seen that the distance error gets bigger by increasing the distance between quadcopter and the center of sensor nodes.

this figure, there are some times where the truth value of localization is less than 0.6 which means the localization has not been carried out properly. This is due to the fact that on those times the quadcopter hits the ground and the propellers are stopped and therefore, no sound is emitted from the quadcopter. As can be seen in this figure, the proposed algorithm could localize the flying object properly in a wide-range area with the size of $(240 \text{ m}, 160 \text{ m}, 80 \text{ m})$.

By using the estimated location (x, y, z) in a period of time, the estimated flying trajectory of the quadcopter can be obtained. The target trajectory is extracted by the GPS module mounted on the quadcopter. In order to depict the trajectory properly, only the XY plane has been shown in Fig. 11. In this figure, the target trajectory is shown by a green dot curve while the blue curve shows the estimated trajectory. As can be seen in this figure, the estimated trajectory is very close to the GPS trajectory which means the proposed algorithm could localize and track the quadcopter properly. As shown in this figure, it can be concluded that greater distance between the sound source and the center of sensor nodes increases the localization error.

As mentioned above, the mean distance error of localization depends directly on the resolution of both the spatial and the angular ROI. Hence, the mean distance errors for different spatial and angular resolutions are brought in Table I. It is worth mentioning that mean absolute error (MAE) in GPS location estimation is about 2.5 m [29], and therefore, the precision of localization is limited to 2.5 m . As can be seen in the table, the minimum of mean distance error is 6.03 m which was achieved when considering the angular resolution of 3° and the spatial resolution of 3 m .

According to the proposed localization algorithm, the distance error is as a result of angular quantization error, mis-calibration error, spatial quantization error and GPS measurement error. Angular quantization error occurs due to the error of direction estimation. In this paper, the mean distance error caused by angular quantization error (ϵ_A) is estimated

TABLE I: Localization mean distance error (meter) considering different spatial and angular resolutions.

Angular resolution θ, ϕ (degree)	spatial ROI resolution (meter)					
	3	5	7	10	15	20
3	6.03	6.30	6.61	7.45	9.61	10.84
4	6.04	6.31	6.69	7.45	9.57	10.98
5	6.17	6.39	7.73	7.46	9.4	10.99
6	6.57	6.79	7.27	7.59	9.36	10.95
7	6.37	6.76	7.2	7.52	9.44	11.17
10	7.77	7.89	8.09	8.53	9.92	11.8
12	6.98	7.26	7.57	8.2	9.38	11.25
15	7.72	7.93	8.34	8.54	10.05	11.47
20	7.54	7.78	8.19	8.47	9.96	11.66
25	8.18	8.3	8.55	9.83	10.63	12.29
30	11.46	11.58	11.6	11.74	12.47	14.15

using Eq. 16 in which A_r denotes the angular resolution for computing the direction of sound source from every node and $\frac{A_r}{4}$ denotes the mean absolute angular error caused by angular quantization. Also in this equation, D represents the average sound source distance from the sensor nodes. Therefore, the mean distance error arising from angular quantization error can be computed using $D \times \sin(\frac{A_r}{4})$. The average number of sensor nodes that receive sound source signals is denoted by N . Since the measurement results of N sensor nodes are fused together, N has inverse effect on the angular error which can be approximated by $\frac{1}{\sqrt{N}}$ based on practical observations.

$$\epsilon_A = \frac{D \times \sin(\frac{A_r}{4})}{\sqrt{N}} \quad (16)$$

Similar to the angular quantization error, mis-calibration of sensor nodes affects the mean distance error. Mis-calibration of the sensor nodes causes error in direction estimation which is a drift angular error in extracted direction. If the MAE of angular mis-calibration of the sensor nodes is represented by A_c , mis-calibration distance error (ϵ_C) can be computed using Eq. 17 which is similar to Eq. 16.

$$\epsilon_C = \frac{D \times \sin(A_c)}{\sqrt{N}} \quad (17)$$

As seen in Table I, spatial quantization error also affects the mean distance error. If the spatial ROI resolution represented by S_r , then the mean spatial quantization distance error (ϵ_S) can be computed using Eq. 18.

$$\epsilon_S = 0.48 \times S_r \quad (18)$$

Since the performance of proposed system is measured by GPS sensor, the GPS measurement error also affects the mean distance error. The MAE of GPS measurement (ϵ_{GPS}) depends on the received signal strength and the number of satellites connected. According to the experiment conditions, the MAE of GPS measurement is estimated about $3m$.

The mentioned four errors ϵ_A , ϵ_C , ϵ_S and ϵ_{GPS} are caused by independent phenomena which affect on the localization precision individually. Moreover, the distance error of the

TABLE II: Ratio of required to available processing resources for implementing the proposed algorithm on the designed sensor node considering different setups.

Setup	#1	#2	#3	#4
Angular resolution for DE	7°	10°	10°	12°
Spatial resolution for SSL	$10m$	$5m$	$10m$	$7m$
Processing resources for DE	88%	43%	43%	30%
Processing resources for SSL	7%	45%	7%	15%
Total processing resources	95%	88%	50%	45%
Mean distance error	$7.52m$	$7.89m$	$8.53m$	$7.57m$

system behaves rather linearly with each of the mentioned errors. Therefore according to the superposition principle, the overall mean distance error is the sum of these errors which can be estimated using Eq. 19. In the practical experiment carried out in this paper, the average sound source distance from the sensor nodes and the angular mis-calibration error are considered as $D = 120m$ and $A_c = 1$, respectively. It should be noted that Eq. 19 is an approximate estimate of the localization mean distance error which can justify the results of Table I with good approximation.

$$\epsilon = \epsilon_A + \epsilon_C + \epsilon_S + \epsilon_{GPS} \quad (19)$$

One of the most important tasks for implementing the proposed algorithm on the designed hardware is the management of processing resources. For instance, the required processing resources of the proposed algorithm when considering the 3° angular resolution and the $3m$ spatial resolution are too large to be implemented on the sensor nodes. On the other hand, the required processing resources are reduced by decreasing spatial and angular resolutions. As mentioned above, there are two main stages in the proposed algorithm: one is the direction estimation and the other is localization. For DE, the amount of required processing resources associates with the second order of angular resolution. For SSL, the amount of required processing resources associates with the third order of spatial resolution. Therefore, the spatial resolution is much more effective on the required processing resources than the angular resolution. Hence, to estimate the location of the sound source in real-time, the spatial resolution, as well as the angular resolution, have to be considered more than a specified value. The ratio of required to available processing resources for different setups on the designed sensor nodes are brought in Table II. For each setup, the amounts of angular and spatial resolutions are considered to be such that the proposed algorithm can be implemented on the designed sensor node and the sensor nodes can estimate the location of the sound source in real-time. The required processing resources for DE and SSL are individually brought in the table. The mean distance error for each setup is also mentioned in this table. It should be noted that there is a tradeoff between the required processing resources and the minimum distance error. The best setup is the one with minimum required processing resources and minimum mean distance error.

As mentioned above, a quadcopter flying in a wide-range outdoor environment has been considered as a sound source

TABLE III: Comparing the precision of different sound source localization applications in wide-range outdoor environment.

Applications	Environment Size (m)	2D or 3D	Node Numbers	Mean Error (m)	Precision (%)
Gunshot localization[34]	60 × 60	2D	9	3	96.5
Elephant localization[35]	600 × 700	2D	4	30	96.7
Wolf localization[36]	6000 × 5000	2D	20	167	97.9
Underwater source localization[6]	1200 × 800	2D	1	112	92.2
Underwater Vehicle localization[5]	80 × 100	2D	2	8	93.8
Underwater Vehicle localization[7]	2500 × 2500	2D	3	80	97.7
Underwater Vehicle localization[8]	120 × 60	2D	2	3	97.9
Quadcopter localization (this paper)	240 × 160 × 80	3D	8	6	98

and localized using the designed sensor nodes. Table III is brought here to compare the results of this paper with those of some other practical sound source localization algorithms. There are some published practical works in wide-range outdoor and also underwater environments which are mentioned in this table. This table consists of different applications for sound source localization: wolf monitoring, elephant localization, gunshot localization, underwater vehicle localization and quadcopter localization. Each of these applications took place in a different environment. The quadcopter was localized in a 3D environment while in the other applications mentioned in this table, the localization was performed in a 2D environment. Moreover, the number of sensor nodes used in these applications was different. According to the applications and their assumed environment, the mean distance errors between the estimated location and the actual location are very different. Hence, mean distance error can not be considered as an evaluation criterion alone. It was therefore necessary to define a common fair criterion for evaluating and comparing the localization precision. The evaluation criterion in this table was defined based on the mean distance error and the maximum possible distance (diameter of the environment). The localization precision of each application was computed using Eq. 20 in which ϵ denotes the mean distance error and d denotes the diameter of test environment. It can be concluded that the performance of the proposed method of localization is comparable with that of the other practical works mentioned in the table.

$$Precision = \frac{d - \epsilon}{d} \quad (20)$$

VI. DISCUSSION AND FUTURE WORKS

In this paper, a sound source localization algorithm with reasonable computational cost and good precision based on fuzzy approach was presented. The results of implementing the proposed algorithm on the designed hardware have shown that the location of a sound source can be accurately estimated in real-time. In order to increase the precision of the sound source localization, we can continue our research based on the following approaches: 1) Increase the angular and spatial resolutions, and 2) Algorithm upgrade for multi-source localization.

By increasing the angular resolution, better precision of sound source localization can be achieved. Since the amount

of required processing resources associates with the second order of angular resolution, increasing the angular resolution makes the system inappropriate for practical applications. In order to increase the angular resolution, it is necessary to increase the scale of time difference of arrival (TDOA), which is equivalent to increasing the size of microphone array. Increasing the size of microphone array is in conflict with portable applications where smaller sensor nodes are more appropriate. Increasing the spatial resolution also causes higher precision of sound source localization. However, the amount of required processing resources associates with the third order of spatial resolution which results in inappropriate for the applications in which the available processing resources are limited. Therefore, although increasing angular and spatial resolutions causes higher precision for sound source localization, but it remarkably increases computational costs.

One of the reasons reduces the precision is the presence of other sound sources next to the target sound source in the environment. The proposed system is capable to localize a single sound source in environment. Therefore, if there are some other sound sources except the target sound source in the environment, the other sound sources are considered as noise which causes an increase in localization error. As a result, by upgrading the algorithm for multi-source localization, higher precision for sound source localization is achieved. Therefore, it seems that localization of multiple sound sources is an appropriate future work which increases the precision. Multiple sound sources localization can be done by separating the various sound sources and then localizing each one individually.

VII. CONCLUSION

In this paper, we have presented a new method in which several sensor nodes were used to address the problem of sound source localization. The paper has presented a design for a sensor node with an array of eight high quality synchronous MEMS microphones. Eight designed sensor nodes were used in this paper. The proposed sound source localization algorithm contains two main stages: DE and SSL. Each sensor node is able to estimate the direction of the sound source using the beamforming method. To remove the installation error of the sensor nodes, the estimated directions were calibrated using a calibration algorithm which leads to extracting the orientation of sensor nodes. To estimate the location of the

sound source, the calibrated directions were fused using fuzzy logic. A practical experiment was carried out to evaluate the proposed sound source localization algorithm. In this experiment, a quadcopter flying in a wide-range area for a specific time period was considered. The location of the flying quadcopter was estimated once in every second. Mean distance error between the estimated and target location was considered as a criterion to evaluate the proposed algorithm. It can be concluded from the experiment that different angular and spatial resolutions leads to different mean distance errors and various computational costs. It is also shown that to localize the sound source in real-time, angular and spatial resolutions have to be restricted to specific values. The experimental results demonstrated the accuracy and feasibility of the proposed approach in sound source localization. The major benefit of the proposed approach is its high precision in sound source localization, robustness and reasonable computational cost, making the proposed algorithm suitable for practical applications.

VIII. ACKNOWLEDGEMENTS

This work was partly funded by EU H2020 grant 824164 HERMES and Spanish grant from the Ministry of Economy and Competitivity TEC2015-63884-C2-1-P (COGNET) (with support from the European Regional Development Fund).

REFERENCES

- [1] A. M. Ali, K. Yao, T. C. Collier, C. E. Taylor, D. T. Blumstein, and L. Girod, "An empirical study of collaborative acoustic source localization," in *2007 6th International Symposium on Information Processing in Sensor Networks*, April 2007, pp. 41–50.
- [2] D. Pavlidi, A. Griffin, M. Puigt, and A. Mouchtaris, "Real-time multiple sound source localization and counting using a circular microphone array," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 10, pp. 2193–2206, Oct 2013.
- [3] C. Zhang, D. Florencio, D. E. Ba, and Z. Zhang, "Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings," *IEEE Transactions on Multimedia*, vol. 10, no. 3, pp. 538–548, April 2008.
- [4] Y. Na, Y. Guo, Q. Fu, and Y. Yan, "An acoustic traffic monitoring system: Design and implementation," in *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, Aug 2015, pp. 119–126.
- [5] M. Bayat, N. Crasta, A. P. Aguiar, and A. M. Pascoal, "Range-based underwater vehicle localization in the presence of unknown ocean currents: Theory and experiments," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 1, pp. 122–139, Jan 2016.
- [6] P. Felisberto, O. Rodriguez, P. Santos, E. Ey, and S. M. Jesus, "Experimental results of underwater cooperative source localization using a single acoustic vector sensor," in *Sensors*, 2013.
- [7] Sheng-Wei Huang, N. Taniguchi, An-Ting Hsiao, Chen-Fen Huang, E. Chen, Chao-Lung Ting, and Jenhwa Guo, "Autonomous underwater vehicle localization using ocean tomography sensor nodes," in *OCEANS 2016 MTS/IEEE Monterey*, Sep. 2016, pp. 1–5.
- [8] N. R. Rypkema, E. M. Fischel, and H. Schmidt, "Closed-loop single-beacon passive acoustic navigation for low-cost autonomous underwater vehicles," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2018, pp. 641–648.
- [9] H. Lim, I. Yoo, Y. Cho, and D. Yook, "Speaker localization in noisy environments using steered response voice power," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 1, pp. 112–118, February 2015.
- [10] A. Lédeczi, A. Nádas, P. Völgyesi, G. Balogh, B. Kusy, J. Sallai, G. Pap, S. Dóra, K. Molnár, M. Maróti, and G. Simon, "Countersniper system for urban warfare," *ACM Trans. Sen. Netw.*, vol. 1, no. 2, pp. 153–177, Nov. 2005. [Online]. Available: <http://doi.acm.org/10.1145/1105688.1105689>
- [11] A. Pourmohammad and S. M. Ahadi, "Real time high accuracy 3-d phat-based sound source localization using a simple 4-microphone arrangement," *IEEE Systems Journal*, vol. 6, no. 3, pp. 455–468, Sept 2012.
- [12] K. D. Frampton, "Acoustic self-localization in a distributed sensor network," *IEEE Sensors Journal*, vol. 6, no. 1, pp. 166–172, Feb 2006.
- [13] M. M. Faraji, S. B. Shouraki, E. Iranmehr, and B. Linares-Barranco, "A fuzzy tdoa approach for sound source localization using a mems microphone array," *under review*, 2019.
- [14] L. Nguyen, J. Valls Miro, and X. Qiu, "Multilevel b-splines-based learning approach for sound source localization," *IEEE Sensors Journal*, vol. 19, no. 10, pp. 3871–3881, May 2019.
- [15] D. Salvati, C. Drioli, and G. L. Foresti, "A low-complexity robust beamforming using diagonal unloading for acoustic source localization," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 609–622, March 2018.
- [16] P. Bergamo, S. Asgari, H. Wang, D. Maniezzo, L. Yip, R. E. Hudson, K. Yao, and D. Estrin, "Collaborative sensor networking towards real-time acoustical beamforming in free-space and limited reverberance," *IEEE Transactions on Mobile Computing*, vol. 3, no. 3, pp. 211–224, July 2004.
- [17] L. Wang and A. Cavallaro, "Acoustic sensing from a multi-rotor drone," *IEEE Sensors Journal*, vol. 18, no. 11, pp. 4570–4582, June 2018.
- [18] L. Battista, E. Schena, G. Schiavone, S. A. Sciuto, and S. Silvestri, "Calibration and uncertainty evaluation using monte carlo method of a simple 2d sound localization system," *IEEE Sensors Journal*, vol. 13, no. 9, pp. 3312–3318, Sep. 2013.
- [19] M. M. Faraji, A. H. Rezaie, and E. Iranmehr, "Real-time ml based algorithm for localizing acoustic source in wsn," in *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, May 2014, pp. 62–66.
- [20] M. M. Faraji, S. B. Shouraki, and E. Iranmehr, "Fuzzy based algorithm for acoustic source localization using array of microphones," in *2017 Iranian Conference on Electrical Engineering (ICEE)*, May 2017, pp. 2102–2105.
- [21] —, "Spiking neural network for sound localization using microphone array," in *2015 23rd Iranian Conference on Electrical Engineering*, May 2015, pp. 1260–1265.
- [22] Y. Li and H. Chen, "Reverberation robust feature extraction for sound source localization using a small-sized microphone array," *IEEE Sensors Journal*, vol. 17, no. 19, pp. 6331–6339, Oct 2017.
- [23] X. Alameda-Pineda and R. Horaud, "A geometric approach to sound source localization from time-delay estimates," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 6, pp. 1082–1095, June 2014.
- [24] V. P. Minotto, C. R. Jung, and B. Lee, "Simultaneous-speaker voice activity detection and localization using mid-fusion of svm and hmms," *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 1032–1044, June 2014.
- [25] Y. Liu, Y. H. Hu, and Q. Pan, "Distributed, robust acoustic source localization in a wireless sensor network," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4350–4359, Aug 2012.
- [26] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S001999586590241X>
- [27] K. LeBlanc and A. Saffioti, "Multirobot object localization: A fuzzy fusion approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 5, pp. 1259–1276, Oct 2009.
- [28] "Low-Noise Microphone with TDM Digital Output," pp. 1–20, 2017. [Online]. Available: <https://www.invensense.com/download-pdf/ics-52000-data-sheet/>
- [29] "Quectel L86: Compact GNSS Module Integrating Patch Antenna," pp. 1–2, 2017. [Online]. Available: <https://www.quectel.com/product/186.htm>
- [30] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. New York, NY, USA: Simon & Schuster, Inc., 1992.
- [31] J. J. Christensen and J. Hald, "Technical review: Beamforming," 2004.
- [32] E. P. Klement and R. Mesiar, "Triangular norms," *Tatra Mountains Mathematical Publications*, vol. 13, 01 1997.
- [33] W. V. Leekwijck and E. E. Kerre, "Defuzzification: criteria and classification," *Fuzzy Sets and Systems*, vol. 108, no. 2, pp. 159–178, 1999. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165011497003370>
- [34] Y. Kim, J. Ahn, and H. Cha, "Locating acoustic events based on large-scale sensor networks," in *Sensors*, 2009.
- [35] C. M. Dissanayake, R. Kotagiri, M. N. Halgamuge, and B. Moran, "Improving accuracy of elephant localization using sound probes,"

- Applied Acoustics*, vol. 129, pp. 92 – 103, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0003682X1730556X>
- [36] M. Papin, J. Pichenot, F. Guérol, and E. Germain, “Acoustic localization at large scales: a promising method for grey wolf monitoring,” *Frontiers in Zoology*, vol. 15, no. 1, p. 11, Apr 2018. [Online]. Available: <https://doi.org/10.1186/s12983-018-0260-2>