

Research Article

Envelope and Wavelet Transform for Sound Localisation at Low Sampling Rates in Wireless Sensor Networks

O. M. Bouzid, G. Y. Tian, J. Neasham, and B. Sharif

School of Electrical, Electronic and Computer Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

Correspondence should be addressed to O. M. Bouzid, o.m.bouzid@ncl.ac.uk

Received 15 November 2011; Revised 12 January 2012; Accepted 15 January 2012

Academic Editor: Shenfang Yuan

Copyright © 2012 O. M. Bouzid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High sampling frequencies in acoustic wireless sensor network (AWSN) are required to achieve precise sound localisation. But they are also mean analysis time and memory intensive (i.e., huge data to be processed and more memory space to be occupied which form a burden on the nodes limited resources). Decreasing sampling rates below Nyquist criterion in acoustic source localisation (ASL) applications requires development of the existing time delay estimation techniques in order to overcome the challenge of low time resolution. This work proposes using envelope and wavelet transform to enhance the resolution of the received signals through the combination of different time-frequency contents. Enhanced signals are processed using cross-correlation in conjunction with a parabolic fit interpolation to calculate the time delay accurately. Experimental results show that using this technique, estimation accuracy was improved by almost a factor of 5 in the case of using 4.8 kHz sampling rate. Such a conclusion is useful for developing precise ASL without the need of any excessive sensor resources, particularly for structural health monitoring applications.

1. Introduction

Integration of acoustic sensors into wireless sensor networks (WSNs) opens up new horizons for developing wired acoustic source localisation (ASL) systems to wireless systems [1]. This involves the utilisation of distributed sensor nodes which are able to realize onboard computation to achieve either distributed or centralized data manipulation. Such integration is adapted to a large variety of applications, including vehicle identification [2], structural health monitoring [3], and military activities [4, 5].

WSNs have been widely used in such applications due to the enormous number of advantages that are highlighted in [6]. This technology is also facing challenges as discussed in [7], where the authors provided a review and a discussion of the various issues associated with WSNs, including bandwidth and computational limitations at the level of sensor node.

High acquisition sampling rate is an important factor in ASL using WSNs. Thus, it still needs to be investigated in order to optimise this valuable technology as will be discussed in this paper. For example, high acquisition sampling rates, which are commonly used in wired ASL systems

[8] based on the Nyquist criterion and required for precise localisation, not only result in a high volume of data but also demand sufficient memory size and a large bandwidth for data transmission. In addition, the heavy traffic load is not applicable in WSNs due to the introduced high latency in data collection and increased power consumption [9].

This means that using high acquisition sampling rates in wireless ASL leads to more hardware complexity, more power consumption, and hence significantly higher production costs. Additionally, for other applications such as structural health monitoring, for example, powerful devices are not always available. This is especially so when the size of sensor node is restricted and power supply is difficult to obtain. Therefore, utilization of low sampling rates in this case will help in solving such problems, since the power consumption is linearly proportional to the sampling rate of an analog-to-digital converter [10]. Recently, there has been interest in the use of low data acquisition sampling rates in WSNs [11], so that low-cost commercial off-the-shelf (COTS) wireless nodes can be implemented without extra hardware.

Nevertheless, lowering the sampling frequency below the Nyquist criterion is detrimental due to information loss, which will produce inaccurate results for sound localisation.

In this case, conventional time-series methods such as cross-correlation (CC) will deliver inaccurate results if they are used directly to estimate the time delay. This is because, in the time domain, the sampling period determines the time resolution which is very low due to the low sampling frequency which means loss in time information. In frequency domain, the frequency contents are violated since we sample below Nyquist criterion.

This problem compels one to combine the time and the frequency domain information which results in time-frequency domain analysis. Such a domain provides information about how the signal content changes with time, thus providing an ideal technique to process and interpret the received signals at low sampling rates. A variety of time-frequency methods have been developed, including the short-time Fourier transform (STFT), Hilbert-Huang transform (HHT), Wigner-Ville distribution (WVD), and wavelet transform (WT). In this work, the last approach is utilized to overcome the challenge of using low sampling rates mentioned early as will be discussed in Section 2.

To the best of our knowledge, there is no previous wireless ASL system yet that uses low sampling rates below Nyquist criterion. Therefore, the goal of this research is to explore the feasibility of using low sampling rates in WSNs to develop low-cost, energy-efficient, and reliable wireless ASL and to achieve a reasonable estimation accuracy of sound location using time-frequency domain analysis, even if the Nyquist rule is violated. More specifically, our contribution in this paper is to

- (1) counteract the impact of using low sampling rates on the estimation accuracy of sound source location by proposing the utilisation of envelope and wavelet transform cross-correlation (EWTCC) in conjunction with parabolic fit interpolation;
- (2) show that as a result of employing EWTCC with parabolic fit interpolation technique low-cost COTS wireless nodes can be used in ASL applications through conducting several wireless ASL measurements and comparing the estimation performance of the proposed EWTCC algorithm by the conventional CC algorithm.

The paper is organized as follows. After explaining the reasons behind lowering sampling rate, Section 2 introduces the proposed method for locating a sound source using a wireless system. Section 3 presents the experimental setup for the wireless ASL system. In Section 4, we discuss and compare the estimation results of sound localisation using both proposed and conventional CC methods. Finally, conclusion is laid out in Section 5.

2. Proposed Localisation Approach at Low Sampling Rates

To the best of our knowledge, there is no specific algorithm yet that is suitable for time delay estimation (TDE) at low sampling rates. However, in [11], it has been shown that for sound source localisation applications low sampling

frequencies can be used in WSNs provided that time-frequency domain and an appropriate analyse technique are utilized. If frequency domain algorithms are used to analyse the aliased version of the received signal, they will be unable to show the dominant spectral component at the original signal frequency. This means that a major portion of frequency contents is lost and any TDE in this domain will lead to inaccurate results. Furthermore, surrounding noise may affect the original signals. Therefore, using the aliased version of the received signal directly in the time or frequency domain will be insufficient and other techniques are, indeed, needed to counteract the effect of violence Nyquist criterion as well as noise effects and to improve the time resolution in order to gain feasible results.

In literature, utilisation of signal envelope instead of amplitudes in the TDE process is one of the preferable methods for accuracy improvement. This has been used in several applications such as ultrasonic ranging measurements [12–15]. The reason behind is to minimize the ambiguity presenting around the onset of signal amplitudes and peak indices if conventional cross correlation (CC) is used.

So far, using the envelope instead of amplitude signal values is an essential step but it is insufficient to establish a robust localisation algorithm at low sampling rates. Further steps are needed to enhance the estimation accuracy. Time-frequency domain algorithm such as WT which uses feature-based contents is a good candidate in this case, because it relies on the analysis of the time and spectral contents of the signal. In addition, WT has a distinctive attribute of utilizing changeable time-frequency windows in the analysis with respect to the conventional windowed cross-correlation method where a constant size window is applied. These significant WT properties help to induce further improvement in the TDE accuracy. Therefore, this kind of approach is selected in this work for extraction of time and frequency contents as well as noise elimination to estimate the time delay amongst received signals.

The proposed technique is a three-stage strategy as shown in Figure 1. In the first stage, the envelopes of the received signals (denoted by $\tilde{x}_i(t)$, $i \in \{1, 2, 3\}$) are extracted using the methods explained in Section 2.1. In the second stage, the WTs of these envelopes are computed. This is satisfied by the utilisation of discrete values of scaling s_j , $j = 1, 2, \dots, N$. For each scaling value, the cross-correlation in conjunction with a parabolic fit interpolation in the wavelet domain is applied to estimate the time delay τ_{12j} . Finally, the average of the computed delays is calculated in order to obtain the final time delay τ_{12} . These steps are explained in more details in Section 2.3.

2.1. Envelope Extraction. As pointed out previously, the utilisation of envelopes rather than the absolute amplitude values of the aliased versions helps to optimise the signal shapes. It also minimizes the ambiguity around peak indices of CC. This means that envelopes present a nonambiguous feature for the sound source localisation.

In the literature, there are several methods, which can be employed to extract the envelopes of captured signals. Actually, any extraction methods can be used here as long as

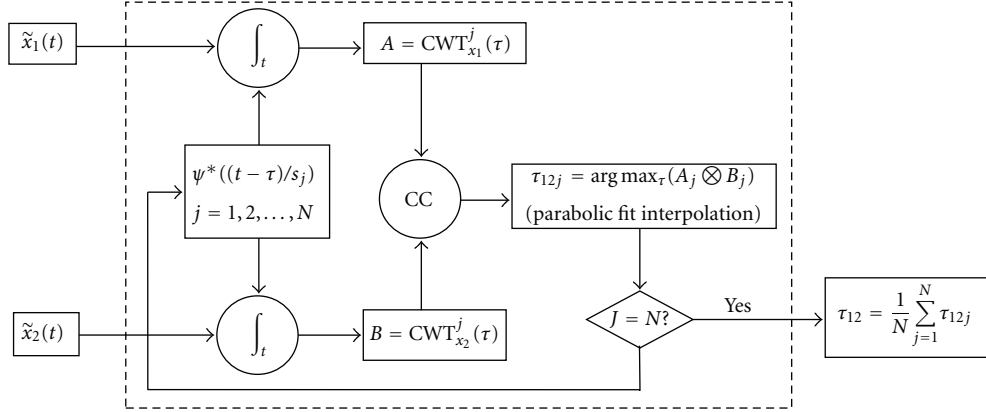


FIGURE 1: Detailed block diagram of the proposed algorithm.

no time delay is introduced due to this operation. Envelopes are usually extracted from bandpass filter outputs by full-wave rectification and lowpass filtering. Another method, which is implemented in this work, is to use the square root of the energies of the original and the Hilbert transformed signals as shown in (1) [16]:

$$\tilde{x}_i(t) = \sqrt{x_i(t)^2 + \hat{x}_i(t)^2}, \quad (1)$$

where $x_i(t)$ is the original signal, $\hat{x}_i(t)$ is the Hilbert transformed signal, $\tilde{x}_i(t)$ is the obtained envelope, and $i \in \{1, 2, 3\}$.

2.2. Wavelet Transform. Both continuous wavelet transform (CWT) and discrete wavelet transform (DWT) have been found that they are effective approaches in many applications including signal processing fields. In this study, we propose to apply WT to counteract the impact of using low sampling rates on the estimation accuracy of a sound source location. In mathematics, WT is defined for a signal $x(t)$ as in (2):

$$\text{WT}_x^\psi(\tau, s) = |s|^{-1/2} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt, \quad (2)$$

where a mother wavelet can be expressed as in (3) [17]:

$$\psi_{(\tau,s)}(t) = |s|^{-1/2} \psi\left(\frac{t-\tau}{s}\right), \quad (3)$$

where $\psi(t)$ is the transforming function (mother wavelet), and $\tau \in \mathbb{R}$ and $s > 0$ are the translation and scale parameters of the mother wavelet, respectively. $1/\sqrt{|s|}$ is an energy normalization factor and “*” denotes the complex conjugate [18]. Equation (2) is also known as the CWT which has the ability to break up a continuous-time function into wavelets through performing an inner production between the signal and a series of son wavelets. These series are generated by stretching and translation of the mother wavelet via controlling s and τ values [19]. Such an operation provides a capability to analyse the signal at different levels of resolutions and to present the processed signal in the time-frequency domain which offers a good time and frequency localisation as explained in the next section.

In addition, thanks to wavelet transform which offers several different valuable mother wavelets that can be employed in the CWT, DWT, and in the signal analysis, including Haar, Meyer, Morlet, Daubechies, Mexican Hat, Gabor, Gaussian, and others. This is, indeed, the strength of this transform which means that based on the signal features we are looking for we can select an optimum mother wavelet to ease our detection of that particular feature. In our proposed technique, we chose the Haar mother wavelet for processing the received signals because its shape is similar to the acoustic signals received by sensor nodes and therefore high correlations between them will result in. On the other hand, noises will be uncorrelated with the Haar wavelet and thus their effects can potentially be reduced in the estimation process.

DWT is another form of WT which involves the use of the dyadic scheme. This is satisfied by the utilisation of discrete values of scaling and translation $s = 2^j, \tau = k2^j, j, k \in \mathbb{Z}$ [17], where \mathbb{Z} denotes the set of integers. In this work, CWT is applied instead of DWT since the last transform is unsuitable for feature extraction [17]. This is because CWT does not require that the wavelet has to satisfy the orthogonality condition which makes it easy for the selection of an appropriate wavelet for feature extraction. Another reason for the utilisation of CWT is that it can be time-invariant which means that the same phase relationship is reserved and no additional time delay is introduced [17, 20, 21].

2.3. Sound Localisation Using Cross-Correlation. The triangular configuration, illustrated in Figure 2, shows the positions of three sensor nodes. They are positioned in a straight line to construct a sensor array with a known geometry. One of the sensor nodes acts as a reference node and is positioned at P_2 . The location of the two other sensor nodes varies between points P_0 and P_4 and accordingly the propagation path differences (PPDs) d_n and d_m also vary. The PPDs are the extra distances that the acoustic signals generated from “S” travel in order to reach the two sensor nodes with respect to the reference node. For simplicity, we assume that these nodes are located at points P_0 and P_4 . Before we describe the

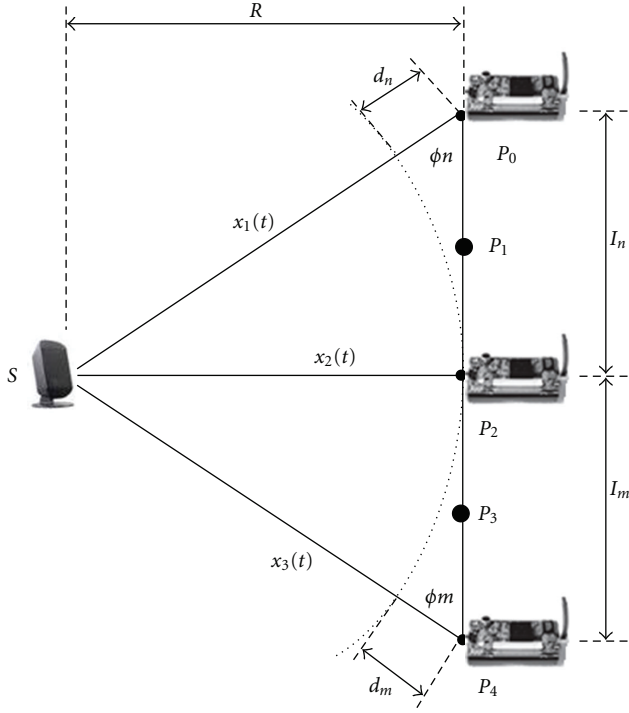


FIGURE 2: Triangular configuration of sensor nodes-sound source for the localization where P_i represents sensor node position and R is the shortest path between sound source “S” and the sensor node located at P_2 .

proposed algorithm to estimate the PPDs, a mathematical model for the acoustic signals received at any microphone is illustrated.

As described in [8, 11] a mathematical model for the acoustic signal captured by any microphone of sensor node $x_i(t)$, $i \in \{1, 2, 3\}$ can be expressed as in (4):

$$x_i(t) = \alpha_i s(t - \tau_i) + \gamma_i(t), \quad (4)$$

where α_i represents an attenuation factor; τ_i is the delay time from the acoustic source $s(t)$ to the i th node; $\gamma_i(t)$ is assumed as a zero mean additive white Gaussian noise with average power $N_0 = 2\sigma^2$, where σ^2 is the noise variance at the i th node.

To estimate the time delay between any two received signals, for instance, $x_1(t)$ and $x_2(t)$, there are many techniques that can be applied to these signals such conventional CC, which is expressed as in (5) [22] and used here for comparison purpose:

$$R_{x_1x_2}(\tau) = \frac{1}{(T - \tau)} \int_{\tau}^T x_1(t)x_2(t - \tau)dt, \quad (5)$$

where T is the observation time interval. The aim of (5) is to examine the coherence between the received signals to estimate the lag at which the CC function has its maximum.

In the proposed algorithm, the CWT is applied to the envelopes of the received signals, for example, $\tilde{x}_1(t)$ and

$\tilde{x}_2(t)$ just before doing CC. Equations (6) and (7) represent the CWT of these envelopes:

$$\text{CWT}_{x_1}^j(\tau) = \int \tilde{x}_1(t)\psi^*\left(\frac{t - \tau}{s_j}\right)dt, \quad (6)$$

$$\text{CWT}_{x_2}^j(\tau) = \int \tilde{x}_2(t)\psi^*\left(\frac{t - \tau}{s_j}\right)dt. \quad (7)$$

As stated in the previous section, varying the s parameter in the mother wavelet in (6) and (7) leads to dilate or compress the signals which allows to searching the similarity in terms of frequency contents between the series of son wavelets and $\tilde{x}_i(t)$ at each scale value: s_j , $j = 1, 2, \dots, N$, where N is the number of variations and τ is assumed to be equal to the sampling period [17]. The process of dilating or compressing the signal via the scale variation allows us to analyse the signal and to compute the wavelet coefficients at different resolutions (multiresolution analysis). The CWT coefficients represent a measure of the cohesion between the signal and the mother wavelet at the current scale. If the frequency components of the signal are corresponding to the current scale of the mother wavelet, then the computed coefficients at this time instant in the time scale are a comparatively large quantity [17]. Two of 2D wavelet coefficients matrices will be generated for both (6) and (7), say A and B , respectively, as shown in Figure 1. Each row in A and B corresponds to the j th wavelet coefficients. The size of these matrices is $(N \times M)$ where M is the length of the processed signal. At each level of resolution the time delay is estimated. As seen in Figure 1 after obtaining the j th wavelet coefficients matrices, the CC algorithm in conjunction with curve-fitting interpolation is applied on the individual rows A_j and B_j and the delay under the j th scale is estimated as in (8):

$$\tau_{12j} = \arg \max_{\tau} (A_j \otimes B_j), \quad (8)$$

where “ \otimes ” denotes conventional cross-correlation. This process is repeated until $j = N$ and then the actual time delay τ_{12} between $x_1(t)$ and $x_2(t)$ can be calculated by taking the average of τ_{12j} as given in (9):

$$\tau_{12} = \frac{1}{N} \sum_{j=1}^N \tau_{12j}. \quad (9)$$

Once the time delays are estimated as shown in the previous paragraph using (9), the PPDs (d_n and d_m shown in Figure 2) can be computed using (10):

$$d_n = c\tau_{12}, \quad d_m = c\tau_{23}, \quad (10)$$

where τ_{12} and τ_{23} are the relative time delays between $x_1(t)$ and $x_2(t)$ as well $x_2(t)$ and $x_3(t)$, respectively. c is the propagation speed of sound in air at room temperature and assumed to be constant in these experiments (340 ms⁻¹). As a result, the acoustic source location can be estimated by applying a triangulation method between the sound source and the three sensor node positions as reported in the following paragraph.

In the following derivation of sound source location in 2D space, we assume that the sensor nodes (one, two, and three) are located at the positions P_0 , P_2 , and P_4 , respectively. Nevertheless, this derivation can be generalized for using any three combinations of sensor locations. From the two triangles SP_0P_2 and SP_4P_2 , we can derive the cosine relations for both angles φ_n and φ_m which are azimuths for sensor nodes one and three, respectively, as in (11) and (12):

$$\cos \varphi_m = \frac{l_m^2 + (R + d_m)^2 - R^2}{2l_m(R + d_m)}, \quad (11)$$

$$\cos \varphi_n = \frac{l_n^2 + (R + d_n)^2 - R^2}{2l_n(R + d_n)}, \quad (12)$$

where l_m and l_n represent the separation distances between sensor nodes as shown in Figure 2 and they are known. R is the shortest path between sound source and reference node. Similarly, from the triangle SP_0P_4 we can develop the expression in (13):

$$\cos(\varphi_m + \varphi_n) = \frac{(l_m + l_n)^2 - (R + d_m)^2 - (R + d_n)^2}{2(R + d_m)(R + d_n)}. \quad (13)$$

Using (11)–(13) it is now possible to calculate, via appropriate substitutions, the three variables R , φ_n , and φ_m as in (14)–(16):

$$R = \frac{l_n(l_m^2 - d_m^2) - l_m(l_n^2 - d_n^2)}{2(d_m l_n + d_n l_m)}, \quad (14)$$

$$\varphi_m = \cos^{-1} \frac{l_m^2 + (R + d_m)^2 - R^2}{2l_m(R + d_m)}, \quad (15)$$

$$\varphi_n = \cos^{-1} \frac{l_n^2 + (R + d_n)^2 - R^2}{2l_n(R + d_n)}. \quad (16)$$

As we can see in (14)–(16), by knowing the variables R , φ_m , and φ_n , we can estimate the sound source location in 2D. Moreover, the propagation path differences d_n and d_m play an important role in estimating these parameters, and the more accurate d_n and d_m are, the better the localisation results become.

As it is known in the classical time delay estimation measurements, the discrete cross-correlation is only calculated at integer indices. This means that it gives an inaccurate estimation if the true delay between two signals is a nonintegral multiple of the sample period. There are several techniques that can be used to optimise this resolution [23]. A common method which is used widely in resolution optimization is to use a parabolic interpolation. This is because the shape of cross-correlation output is similar to the Gaussian curve where the position of the peak is located at the center of this curve. Theoretically, fitting a parabola requires at least three points as shown in Figure 3: the maximum peak of correlation coefficients, its preceding, and subsequent neighbors. The blue dotted curve represents the fitted curve

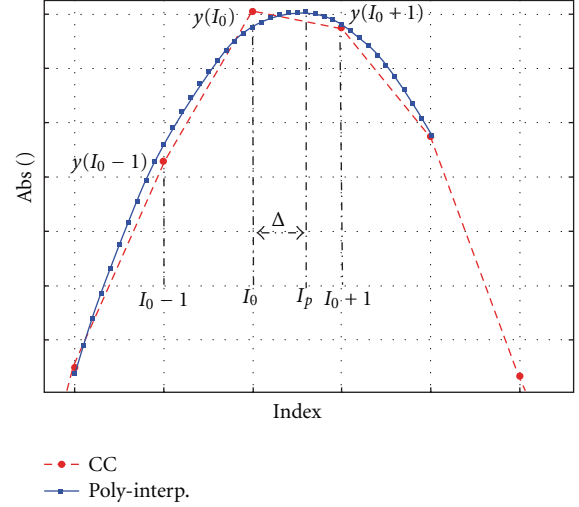


FIGURE 3: Curve-fitting interpolation using three points.

to the cross-correlation output (red dashed curve). These three points are needed to calculate the coefficients a , b , and c in (17) which represents the applied parabola [23]:

$$y = ax^2 + bx + c. \quad (17)$$

To use this polynomial in the fitting process, we need to first calculate the coefficients a , b , and c . After that, by taking the derivative of (17) which equals zero at the maximum peak, we can compute the interpolated peak, I_p , as illustrated in (18):

$$I_p = -\frac{b}{2a}. \quad (18)$$

A series of experiments have been conducted concentrating on the estimation of the propagation path differences to show the performance of the proposed approach.

Based on the used sample frequency (4807 Hz), the minimum distance resolution is 7.07 cm. To improve this resolution, we apply the parabolic fit interpolation on the output of the EWTCC. This results in an improvement in this distance resolution from 7.07 to 1.50 cm, which is almost five times better. Such a resolution improvement will contribute to the estimation accuracy of sound source location using WSNs at low sample rates as illustrated in the next section.

3. Experimental Work

The wireless acoustic source localisation system used in this work is depicted in Figure 4. The system was employed to study the utilisation of a single-hop WSN for sound source localisation at low sampling frequencies. Three acoustic sensor nodes were placed in a straight line. The sensor nodes are (MICAz) motes equipped with the sensor board (MTS310), which has different sensor modalities. The nodes sense simultaneously the omnidirectional microphone sensor modality of the sensor boards and send the data to a base station. All the sensor nodes communicate with the

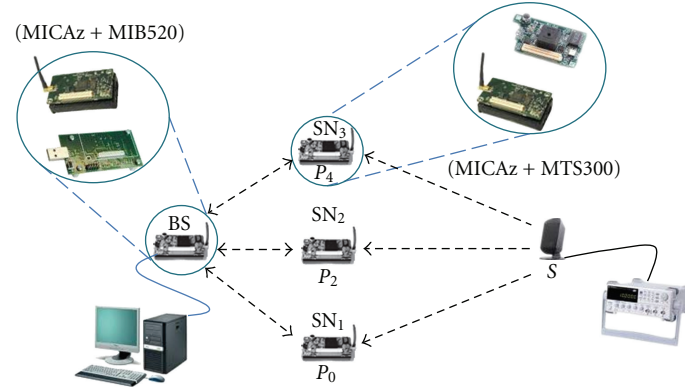


FIGURE 4: Three sensor nodes (SN_1 , SN_2 , and SN_3) positioned at locations P_0 , P_2 , and P_4 , respectively, are used for acoustic source localization.

base station via an RF interface. The base station is plugged into the gateway board: MIB520, and it is used to forward the received signals to a PC where they are processed. The sensor nodes are programmed under TinyOS (TinyOS is a tiny operating system and has been widely used for WSNs design) environment. A Listen application is used to eavesdrop on messages sent over the mote radios and the data received are saved in a hexadecimal format. The received data are processed off-line using MATLAB. Time delays and the propagation path differences are estimated from acoustic signals captured using the approaches explained in Section 2.

Since we sample much below the Nyquist criterion and our proposed algorithm for TDE is based on the shapes contents of received signals, we have assumed that the acoustic signal which will be used as a test signal in the experimental work is a narrow bandwidth and not periodic signal because it is the type of signal we expect from a real scenario. This means that the generated acoustic signals has a finite pulse duration and the repetition period of this pulse is greater than the sampling duration in order to have one spike vanishes before the next spike is started. The length of the sampling duration also should be long enough to be admitted in order to collect a sufficiently amount of samples that represents a complete pulse.

Based on the aforementioned conditions an acoustic pulse test signal is used and simulated using a function generator through generating a tone burst of 50-sinusoidal cycles of frequency 10 kHz. The test signal is played through a PC speaker and such values are selected experimentally in order to generate a reasonable pulse shape for the conducted experiments. The generated acoustic signals are acquired at a 4807 Hz sampling rate for a sampling duration of 0.25 s. The 4807 Hz is almost the lowest sampling rate that MICAz mote can achieve using the hardware event handler (HEH) mode while in [11] it has been shown that sampling rates below this sampling rate can be used. An important condition here for reducing the sampling rate is that the samples should contain enough samples from the acoustic signal acquired in order to extract the envelope of the signal which should be sufficient to use in the time delay estimation process.

The experiments were conducted in an ordinary indoor laboratory environment which has objects, such as tables, PCs, and equipment. Street traffic and people talking contribute to the background noise where the experiments were being conducted. They were conducted as follows: a base station broadcasts a start sample command. Once sensor nodes receive it, they start sensing until the buffer becomes full (1200 data points). At this point, each node starts to send the acquired data back to the base station. To avoid data collisions nodes will send their own packets in sequence to the base station. In the future work, it is planned to process the received signals locally in the sensor node so that the advantage of utilizing low sampling rates will be obviously. The received data will be processed as shown in the next section.

4. Results and Discussion

In the conducted experiments, for the evaluation study the three sensor nodes were arranged in a straight line at different positions (P_0 - P_4) as shown in Figure 4. Once the sensor nodes receive simultaneously a start sample command, they commence to acquire the generated acoustic signal as explained in Section 3.

Before presenting and discussing the results of this work, it is significant important to mention that sound source localisation using WSNs depends highly upon time synchronization among the sensors. However, from previous work in the area [24] and our preliminary experimental results, it is noticeable that the use of time-synchronized sensor nodes based on a global time does not guarantee that the acquired acoustic signals (i.e., sensing operation) are perfectly synchronized with each other. The adopted approach in performing a data acquisition operation also has a significant impact on achieving synchronized data acquisition operations. In general, nondeterministic (if the execution time of the same code varies at each repeated execution) operating systems have the disadvantage that they do not allow the user to control the execution process of their measurement process (i.e., setting priorities to the measurement steps). Such a property makes the execution of the data acquisition operations start at different time instants and also it does not

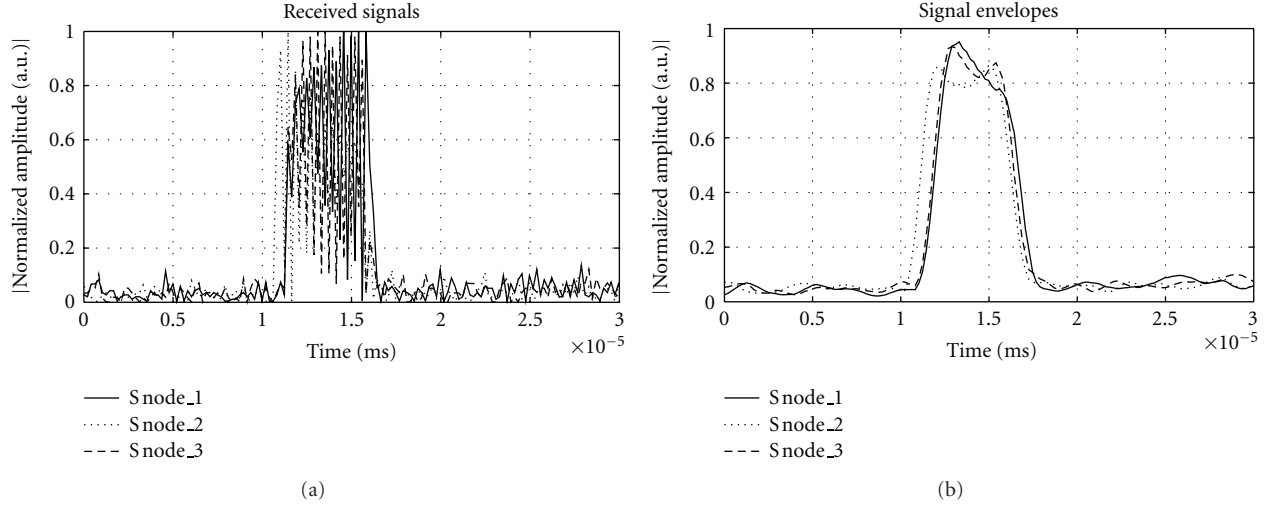


FIGURE 5: Received signals (zoom-in): (a) captured signals, (b) signal envelopes.

take the same amount of time for every execution, especially if it is performed on different microcontrollers, as in the case of WSNs.

TinyOS, which is used in this work as well as others [24], has two modes of execution threads: tasks and hardware event handler (HEH) [25]. The first mode has a nondeterministic nature, which introduces an unpredictable waiting time during the acquisition operation due to the TinyOS scheduler as it executes posted tasks. This results in unequal intervals which makes acoustic sensing tasks at all sensor nodes unsynchronized. To counteract this problem, the HEH mode is proposed to realize a synchronized data sensing amongst sensor nodes because it is a deterministic mode and asynchronous commands are executed immediately.

To verify this, sensor nodes were located at the same point where real PPD values are zeros. This configuration represents the extreme case and helps to test the performance of both modes (tasks and HEH) as it requires a high time resolution for accurate TDE from amongst the received signals. The experiments were repeated 15 times, and in each time, the PPDs were estimated through applying EWTCC algorithm to acquired signals as described in Section 2. The results of this test show clearly that the HEH mode exhibits much better performance than tasks mode. The RMS error of HEH mode is 1.7 cm whereas it is 24.11 cm for tasks mode case.

From the previous discussion, we can conclude that the tasks mode is not applicable in achieving a synchronized sensing process amongst all nodes. In contrast, the HEH mode results show much less synchronization errors. Minimizing these errors results in an improvement in the estimated sound source location. This is because such errors are accumulated to the real-time delays among received signal. Therefore, the HEH mode is chosen for realizing a synchronized data acquisition operation and designing a wireless sound localisation system.

Figure 5(a) shows the acoustic signals (zoom-in) captured using sensor nodes 1, 2, and 3 as they were positioned

at P_0 , P_2 , and P_4 in Figure 2, respectively. In contrast, Figure 5(b), depicts the signals envelopes extracted from these signals. In both figures, the solid curves represent the signals received at sensor node one and the dashed curves show the signals received at sensor node three. While the dotted curves illustrate the signals received at sensor node two (reference). As we can see in Figure 5(b), the extracted envelopes are much clearer than the original signals and therefore feeding these envelopes, instead of the original signals, to the Wavelet transform is an important step towards improving the localisation accuracy. In the conducted experiments, the distance between the reference node and the sound source represents the shortest path and that why the received signal from this sensor node appears first in Figure 5, whereas the other two sensor nodes were located equally apart from the reference node. Therefore, their received signals appear second and at the same time instance.

Figure 6 illustrates (zoom-in) the results of applying the CC and EWTCC with and without curve-fitting interpolation on envelopes shown in Figure 5(b). Such results illustrate the functionality of the proposed approach and give an example about how it improves the spatial resolution of the sound source localisation. The advantage of using the EWTCC here over the traditional CC method is to enhance the signal-to-noise ratio (SNR) due to the high correlation between these envelopes and used the Haar mother wavelet as well as to sharp the output of CC which makes the identification of the final (most accurate) index much easier when curve-fitting interpolation is utilised.

The experiments were also conducted for different values for d_m and d_n (i.e., positions of sensor nodes one and three were varied between P_0 and P_4). The results of these experiments are reported in Table 1. This table summarises the estimation results obtained by the developed wireless localisation system using both conventional CC and our proposed EWTCC with and without curve-fitting interpolation approaches in d_n and d_m estimations. It summarises the averages and the standard deviations of test results

TABLE 1: Averages (AVGs) and standard deviations (STDs) of estimated propagation path differences (PPDs) for 15 experiments.

Configuration of SN_1 SN_2 SN_3	Actual PPDs (cm)	CC		Estimated PPDs (cm)		EWTCC- Fitting	
		AVG	STD	AVG	STD	AVG	STD
P_2 P_2 P_2	d_m	00.00	6.14	1.83	1.17	1.15	0.81
	d_n	00.00	2.56	1.96	1.87	1.27	1.06
P_0 P_2 P_3	d_m	15.00	12.26	14.10	3.53	15.19	2.64
	d_n	36.00	26.41	37.11	3.89	36.49	3.71
P_1 P_2 P_3	d_m	21.00	23.11	22.95	5.02	21.40	4.12
	d_n	15.00	12.26	16.55	5.69	15.65	4.53

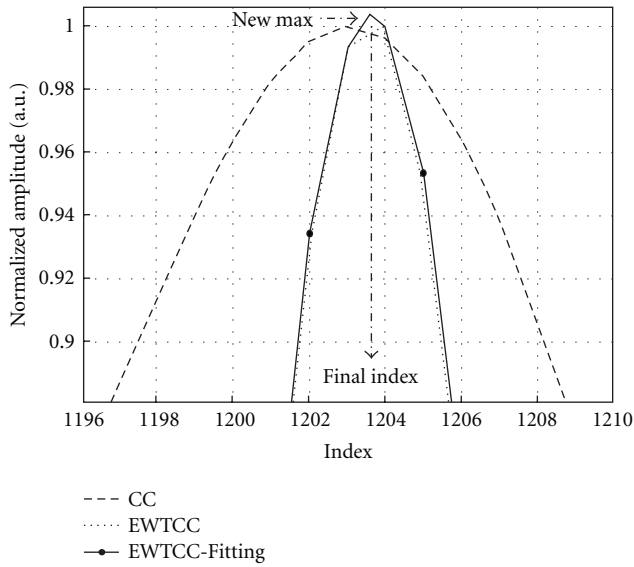


FIGURE 6: Output (Zoom-In) of CC and EWTCC with and without parabolic fit interpolation.

gained from 15 experiments in order to show that replicated measurements can provide closely similar results. It is apparent from this table that results of applying EWTCC with curve-fitting interpolation are much more correlated to the real values of d_n and d_m than the other results. A good example of this is the case where $d_n = 15$ cm. The average of the estimated result using EWTCC with fitting is 15.65 cm, while it is 12.26 cm using CC and 16.55 using EWTCC without fitting. Again, the variation of the EWTCC with fitting result is 4.53 compared to the other two cases. This means that errors in estimation of d_n and d_m using EWTCC with curve-fitting interpolation are much less than errors in the CC method due to the multiresolution analysis property of CWT and curve interpolation. Consequently, the use of such method in estimation of PPDs exhibits better performance than employing the CC method as shown in Figure 7.

Figure 7 illustrates the RMS computed errors for the three configurations shown in Table 1. As seen from this figure, the maximum RMS errors in estimation of d_n and d_m using EWTCC with curve-fitting interpolation is 1.70 cm while it is 2.68 cm using just EWTCC and 9.97 cm using

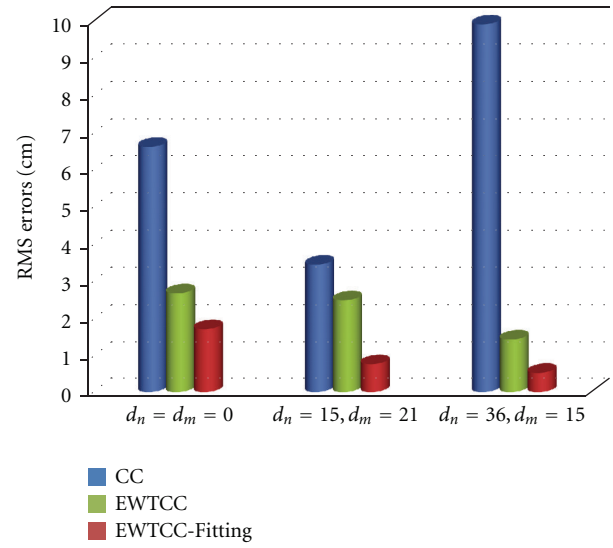


FIGURE 7: RMS errors for CC, EWTCC, and EWTCC-Fitting algorithms.

CC. In addition, the trend of the RMS errors using the proposed method shows that RMS errors decreases with moving away from the worst case scenario, whereas, they are randomly for the CC algorithm case. Such an enhancement in the estimation accuracy of sound localisation correlated to two reasons: (1) employing envelopes of acquired signals reduces the ambiguity present around peak indices of CC; (2) processing these envelopes in time-frequency domain using WT integrates both time and spectral contents in the estimation-process. Thus, EWTCC algorithm in conjunction with curve-fitting interpolation is able to achieve a sufficient level of estimation accuracy for wireless ASL at low sampling rates compared to the CC method.

5. Conclusion

Envelope and wavelet transform cross-correlation, EWTCC, in conjunction with a parabolic fit interpolation method is proposed for wireless ASL employing low sampling rates. The new technique, in comparison to conventional CC algorithm, offers a multiresolution analysis domain which shows a potential performance in counteracting the ambiguous

peaks due to the low time resolution. The proposed approach also enhances the spatial resolution of the localisation process from 7.07 cm to 1.50 cm. Such results of the conducted experiments show the consistence and low errors for the case study; further evaluation work can be done for large-scale measurement, including complex geometrical and local data processing scenarios as well as optimisation of the estimation process through selection of the best scale values which delivers the best resolution. In addition, the proposed HEH mode realizes a synchronized data acquisition operation for all sensor nodes in the TinyOS-based WSNs. It must be emphasized here that such conclusions can open up new horizons for the development of efficient low-cost, reliable wireless ASL systems based on low-cost COTS sensor nodes without the need to support excessive sensor resources, as low sampling rates not only contribute to a cost reduction but also minimize power consumption and extend the lifetime of sensor nodes which allows having the processing operations in real time.

Acknowledgment

The authors would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for funding the work through the project of Instrument for Soundscape Recognition, Identification and Evaluation (EP/E008275/2).

References

- [1] J. C. Chen, L. Yip, J. Elson et al., "Coherent acoustic array processing and localization on wireless sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1154–1161, 2003.
- [2] M. F. Duarte and Y. H. Hu, "Vehicle classification in distributed sensor networks," *Journal of Parallel and Distributed Computing*, vol. 64, no. 7, pp. 826–838, 2004.
- [3] A. Lédeczi, T. Hay, P. Völgyesi, D. R. Hay, A. Nádas, and S. Jayaraman, "Wireless acoustic emission sensor network for structural monitoring," *IEEE Sensors Journal*, vol. 9, no. 11, pp. 1370–1377, 2009.
- [4] M. Winkler, K.-D. Tuchs, K. Hughes, and G. Barclay, "Theoretical and Practical aspects of military wireless sensor networks," *Journal of Telecommunications and Information Technology*, vol. 2, pp. 37–45, 2008.
- [5] J. Sallai, W. Hedgcock, P. Völgyesi, A. Nadas, G. Balogh, and A. Ledeczi, "Weapon classification and shooter localization using distributed multichannel acoustic sensors," *Journal of Systems Architecture*, vol. 57, no. 10, pp. 869–885, 2011.
- [6] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.
- [7] A. Jangra, Swati, Richa, and Priyanka, "Wireless sensor network (WSN): architectural design issues and challenges," *IJCSE: International Journal on Computer Science and Engineering*, vol. 2, no. 9, pp. 3089–3094, 2010.
- [8] H. Atmoko, D. C. Tan, G. Y. Tian, and B. Fazenda, "Accurate sound source localization in a reverberant environment using multiple acoustic sensors," *Measurement Science and Technology*, vol. 19, no. 2, Article ID 024003, 2008.
- [9] C. T. Cheng, C. K. Tse, and F. C. M. Lau, "An efficient data collecting network structure in wireless sensor networks," in *Proceedings of the International Conference on Communications, Circuits and Systems (ICCCAS '08)*, pp. 444–447, May 2008.
- [10] K. Philips, $\Sigma\Delta$ A/D conversion for signal conditioning, Ph.D. thesis, 2005.
- [11] O. M. Bouzid, G. Y. Tian, J. Neasham, and B. Sharif, "Investigation of sampling frequency requirements for acoustic sourcelocalisation using wireless sensor networks," *Applied Acoustics*. In press.
- [12] R. Raya, A. Frizera, R. Ceres, L. Calderón, and E. Rocon, "Design and evaluation of a fast model-based algorithm for ultrasonic range measurements," *Sensors and Actuators, A*, vol. 148, no. 1, pp. 335–341, 2008.
- [13] F. E. Gueuning, M. Varlan, C. E. Eugène, and P. Dupuis, "Accurate distance measurement by an autonomous ultrasonic system combining time-of-flight and phase-shift methods," *IEEE Transactions on Instrumentation and Measurement*, vol. 46, no. 6, pp. 1236–1240, 1997.
- [14] A. Egaña, F. Seco, and R. Ceres, "Processing of ultrasonic echo envelopes for object location with nearby receivers," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 12, pp. 2751–2755, 2008.
- [15] X. L. Xu, A. H. Tewfik, and J. F. Greenleaf, "Time delay estimation using wavelet transform for pulsed-wave ultrasound," *Annals of Biomedical Engineering*, vol. 23, no. 5, pp. 612–621, 1995.
- [16] B. S. Lazarov, R. Matzen, and Y. Elesin, "Topology optimization of pulse shaping filters using the Hilbert transform envelope extraction," *Structural and Multidisciplinary Optimization*, vol. 44, no. 3, pp. 409–419, 2011.
- [17] J. Lin and A. Zhang, "Fault feature separation using wavelet-ICA filter," *NDT and E International*, vol. 38, no. 6, pp. 421–427, 2005.
- [18] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 569–586, 1992.
- [19] K. C. Ho and Y. T. Chan, "Optimum discrete wavelet scaling and its application to delay and doppler estimation," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2285–2290, 1998.
- [20] J. Lin and L. Qu, "Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis," *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135–148, 2000.
- [21] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.
- [22] J. P. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 6, pp. 998–1003, 1982.
- [23] N. S. M. Tamim and F. Ghani, "Techniques for optimization in time delay estimation from cross correlation function," *IJET/IJENS: International Journal of Engineering & Technology*, vol. 10, no. 2, pp. 39–68, 2010.
- [24] T. Nagayama, S. H. Sim, Y. Miyamori, and B. F. Spencer, "Issues in structural health monitoring employing smart sensors," *Smart Structures and Systems*, vol. 3, no. 3, pp. 299–329, 2007.
- [25] P. Levis and D. Gay, *TinyOS Programming*, Cambridge University Press, Cambridge, UK, 2009.

