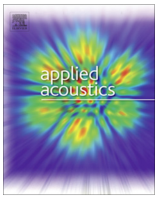
Applied Acoustics 113 (2016) 162–169

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust



An evaluation of low-power microphone array sound source localization for deforestation detection

Lucian Petrica

University Politehnica of Bucharest, 313 Splaiul Independentei Street, 060042 Bucharest, Romania
National Institute for Research and Development in Microtechnologies, 126A Erou Iancu Nicolae Street, 077190 Bucharest, Romania

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form 21 June 2016
Accepted 23 June 2016
Available online 30 June 2016

Keywords:
Sound source localization
Beam forming
FPGA

ABSTRACT

Illegal deforestation is a worldwide problem which may be alleviated through technological means of deforestation monitoring, e.g. wireless sensor networks capable of identifying chain-saw noise, performing sound source localization (SSL), and alerting the authorities to the location of the illegal deforestation activity. In this paper we evaluate the feasibility of performing SSL on low-power, energy-constrained, microphone-array-equipped sensor nodes (SNs) with the Delay-and-Sum (DS) beam-forming algorithm. Our work is the first application of this technique for chain-saw noise. We evaluate array configurations of 4, 8, and 16 microphones, and a multitude of DS algorithm configurations, utilizing chain-saw recordings from the ESC dataset, which is available online. We implement the DS algorithm as a digital circuit in a Xilinx Spartan-6 FPGA and analyze its energy consumption. Our analysis indicates that accurate chain-saw localization can be achieved with much simpler microphone arrays and DS configurations compared to previous work. Furthermore, adding FPGA-based SSL capability to the SN increases the energy consumption by less than 10%, compared to a baseline SN capable only of chain-saw identification through spectral analysis executed in software on the SN microcontroller.

1. Introduction

Up to 40% of the deforestation performed worldwide is illegal in nature [1], causing an estimated 12% of total CO_2 emissions [2]. Wireless sensor networks (WSNs) with acoustic sensing capabilities have been proposed as an efficient deforestation monitoring solution, capable of detecting chain-saw noise in sensitive forest areas and alerting authorities. A WSN for deforestation monitoring consists of a number of microphone equipped sensing nodes (SNs), capable of analyzing audio data individually or cooperatively in order to determine whether deforestation activities are taking place in the monitored area, e.g. chain-saw cutting [3,4]. For chain-saw identification in particular, several algorithms and their WSN implementations have been proposed and evaluated in [5,6].

This type of monitoring is useful in some situations where logging is completely prohibited, e.g., in natural reservations. A more complex but also more realistic use-case for a WSN is when cutting is permitted in certain areas of the forest but prohibited in others such as steep slopes or riverbanks. Licensed loggers sometimes ignore such restrictions, looking to maximize the profit from an allocated license. This scenario is very difficult for WSNs proposed in previous work because the localization, i.e., finding the approximate position of the deforestation activity, is required in addition to detection. The simplest form of localization is identifying the sound source direction. To date, there is no realistic evaluation of SSL for WSNs in the scientific literature.

Fig. 1 illustrates four sensor nodes of a WSN and a triangular area of restricted logging, e.g., a steep slope, surrounded by an area where logging is permitted. If as in Fig. 1a logging takes place in the restricted area, the sound propagates outward and reaches the SNs. Without localization capability, as in Fig. 1b, the WSN cannot determine if the logging activity is legal or not, only that there is a higher probability of the sound originating somewhere within the polygon enclosed by the SNs. The range of each SN, i.e., the maximum distance at which it is able to detect logging noise, is a factor of the acoustic propagation environment [7,8], which introduces further uncertainty. SSL-enabled SNs, capable of identifying the direction of the incoming sound, provide a clear advantage, as illustrated in Fig. 1c. Each SN determines the direction of the incoming sound relative to itself and communicates this information to other SNs through radio. By superimposing direction estimations from two or more SNs, the exact source of the sound may be identified: the sound source is located where the individual direction estimations intersect. An alarm may be raised if the source is within the restricted area. A sound source may also be localized

---

E-mail address: lucian.petrica@upb.ro

http://dx.doi.org/10.1016/j.apacoust.2016.06.022
0003-682X/© 2016 Elsevier Ltd. All rights reserved.

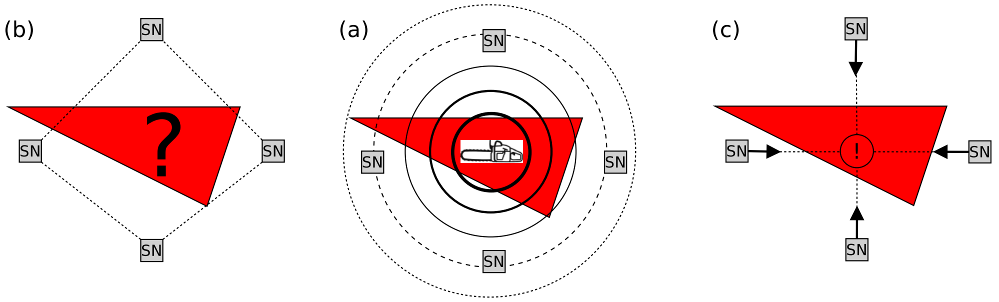


Fig. 1. Discrimination between legal and illegal deforestation by a SSL-enabled WSN.

outside of the polygon enclosed by the sensor nodes, as long as it is within the sensing range of at least two sensor nodes, allowing for the previously mentioned intersection to take place.

Sound source localization has been achieved for single-microphone SNs in previous work through the use of distributed time difference of arrival (TDOA) [6,9]. In TDOA, the WSN functions as a microphone array. To perform the localization, nodes sample the sound field at a pre-determined moment in time. Because the sampling is simultaneous, differences in audio signal phase at the sensing nodes is caused by differences in relative position of the sound source to the respective SNs. If the geometry of the WSN is known a priori, then the individual measurements may be aggregated to determine the position of the sound source relative to the WSN, through a beam-forming algorithm, e.g., Steered Response Power – PHAse Transform (SRP-PHAT) [10]. As SRP-PHAT requires all the acoustic data captured by each SN, the data must be transported from each SN to a central processing location. Depending on the type of sensor, communication energy costs for WSN nodes is 9 to 22 times larger than computation and sensing energy costs [11,12], therefore this centralized beam-forming exerts a large energy cost on the WSN.

A different, distributed TDOA approach to sound source localization has been explored in [13] utilizing a 52-micro-phone multiple-ring circular array and a FPGA on each SN to execute a Delay-and-Sum (DS) [14,15] beam-forming algorithm for sound direction estimation. DS has been selected for its low computational requirements and because it is implementable in Field Programmable Gate Array (FPGA) hardware. The multi-microphone data gathering and processing in TDOA is above the capabilities of a microprocessor and a FPGA must be utilized for signal acquisition and processing. Therefore, at the cost of increased SN complexity, sound source localization may be performed without significant inter-SN communication. DS-based SSL has been evaluated in [13] on monochromatic sounds and found to provide good localization accuracy for high-pitched sounds.

While work in [13] is conclusive with regard to the general characteristics of microphone array SSL for monochromatic sounds, it remains an open question the extent to which the DS algorithm is effective for chain-saw and tree cutting sound, which is inherently noisy and complex. We evaluate its effectiveness for chain-saw localization, utilizing chain-saw recordings from the ESC [16] online dataset, and perform extensive design space exploration in order to identify the best algorithm parameter configurations, with regard to SSL accuracy.

Furthermore, we demonstrate that microphone array chain-saw localization is possible and effective even in hardware configurations much reduced compared to the 52-microphone array utilized in [13]. Less microphones are expected to reduce the cost of a SN, and the energy consumption during SN operation, thus extending WSN life. Finally, we perform an exploration of the FPGA design parameters in order to determine how factors such as FFT size and the number of audio samples analyzed affect the dimensions of the FPGA circuit and its power dissipation. While the authors of [13] claim a FPGA implementation of their algorithm, no specifics are described in their work, e.g. which is the smallest FPGA device which may implement DS, and no power and energy analysis is performed. Our paper provides this missing information. We evaluate SSL from the perspective of energy consumption, which is the most important WSN metric, and demonstrate that microphone array SSL executed on FPGA consumes less than 2 mJ per localization, and increases SN energy consumption by under 10% compared to a baseline SN which performs only chain-saw identification in software on the SN microcontroller. Together, these contributions demonstrate that adding microphone array SSL capability to WSNs is not only possible, but also practical.

2. Delay-and-Sum SSL implementation

The WSN proposed in Fig. 1 consists of a number of sensor nodes which communicate sound localization information through radio, e.g. WiFi. Previous implementations of deforestation detection WSNs in [17] utilize Mica [18] sensor nodes which do not have sound direction localization capability, which is required by our WSN concept. To add this capability to each Mica sensor node, we designed an add-on printed circuit board which includes a circular microphone array and FPGA. While we did not fabricate the add-on board, the design is sufficiently detailed to enable us to evaluate its performance through simulation.

2.1. Microphone array

Our design utilizes a single-ring uniform circular microphone array which may be populated with 4, 8 or 16 microphones. The uniform circular array layout has the advantage that its estimation accuracy does not depend on the direction of arrival (DOA) of incoming sound, unlike a linear array [19]. The number of microphones is denoted N_M. The radius of the array is 10 cm as in [13], a reasonable size for a SN. Since in a WSN application, most chain-saw localizations will occur at ranges of 10–50 m, we will consider that the sound source is in the far field.

Fig. 2 illustrates an 8-microphone array. The array plane is horizontal, and we assume sound sources occur in the array plane. The principal directions (PDs) of the array are radial directions which connect the center of the array with each of the array microphones. Because the array is symmetric, any PD may be chosen as the root *PD*0, and all other PDs are named according to the counterclockwise angle between themselves and PD0. The principal directions are separated by the step angle θ_S, the value of which is expressed in Eq. (1) in degrees, while the angle between *PD*0 and the DOA is denoted θ_I.

$$\theta_S = \frac{360}{N_M} \tag{1}$$