Seismic Detection of Elephant Footsteps

Daniel Goderik†, Albin Westlund†, Gustav Zetterqvist⋆, Fredrik Gustafsson, Gustaf Hendeby
Division of Automatic Control, Linköping University, SE-581 83 Linköping, Sweden
Email: {dango893, albwe662}@student.liu.se,
{gustav.zetterqvist, fredrik.gustafsson, gustaf.hendeby}@liu.se

***Abstract*—As human settlement expands into the natural habitats of wild animals, the conflicts between humans and wildlife increases. The human-elephant conflict causes a tremendous amount of damage, often to poor villages close to the savannah. In this paper, we continue our earlier reported research on a geophone network aimed for elephant localisation by focusing on the detection challenge. We have now collected larger sets of seismic data with footsteps from both elephants and other big animals including humans. To detect the footsteps, a method is developed that analyses features of the geophone signal, which are then compared to those of an elephant footstep. The method detects 54 % of the footsteps and has a classification accuracy of 89 %. Subsequently, the detected elephant footstep is used to calculate the *direction of arrival* (DOA) angle using a delay-and-sum beamformer. The direction to an elephant is estimated with good precision on distances ranging from 8 to 30 meters. This research, not only, showcases a practical solution for mitigating human-elephant conflicts, but also underscores the potential of seismic technology in wildlife management and conservation efforts.**

***Index Terms*—Elephants, Detection, Direction of Arrival, Geophone network**

I. Introduction

In many parts of the world, the human-wildlife conflict poses a great problem for both humans and wildlife. As human settlement expands, it expands into the natural habitats of wild animals. This causes the human-wildlife conflict to become more and more prevalent. In particular, elephants have been shown to cause more damage to humans than any other herbivore species [1]. Elephants can cause a tremendous amount of damage to villages struck with poverty by damaging infrastructure, consuming farmers crops, or even killing people. This inevitably causes irritation among humans and often leads to killing of the elephants, in lack of better methods. This is one of the biggest issues in elephant conservation efforts today [2].

Traditional methods for deterring elephants, such as electric fences, are expensive to build, have large maintenance costs and are, in general, fairly unreliable [3]. There is a need for a modern, inexpensive and reliable method to keep track of elephants near human habitats to mitigate the human-elephant conflict.

An innovative approach is to utilise the distinctive seismic signatures generated by different animals' footsteps. With knowledge of these seismic impacts, it would be possible to differentiate an elephant's footstep from those of other species. By measuring seismic signals from the ground with geophones, an elephant footstep can be detected from distances that allow park rangers to intervene before the elephants reach the villages [4]. Recent research has shown that seismic signals from elephant footsteps can be perceived up to 155 meters away [5]. In this paper, a device consisting of three geophones is constructed. The geophones measure seismic signals from the ground to detect and localise elephants, as illustrated in Fig. 1.

Fig. 1: Illustration of the intended purpose of the paper – detect and localise an elephant using a geophone network.

There are different kinds of seismic waves that can propagate from a source. The two main types of waves are *body waves* and *surface waves*. Body waves are able to propagate deep into the earth, whereas surface waves travel predominantly along the surface. For this article, surface waves are most relevant, since they are the ones that can be measured by geophones [6]. Based on previous research, elephant footsteps usually consist of a type of surface waves called Rayleigh waves, which rarely exceeds 200 Hz [7], [8].

Previous studies have indicated that elephant footstep signals mainly have a frequency content between 10 and 20 Hz, with the majority of the signal energy concentrated below 40 Hz [9]. Building upon these findings, this study draws inspiration from existing methodologies for detecting elephant footsteps through frequency analysis [4].

While the concept of using geophones for footstep detection is not novel, previous research have primarily targeted human footsteps for security or military purposes. For instance, in [10], an algorithm that detects human footsteps is developed.

The authors marked with a † have contributed equally to this work. The corresponding author is marked with ⋆. G. Zetterqvist has received funding from ELLIIT. This work was partially funded by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Their approach relies on kurtosis to detect footsteps, and examines the cadence, *i.e.*, the time between footsteps, to determine if the source is human. Similarly, in [11], kurtosis is employed for footstep detection, indicating its popularity, although it can not differentiate between different animals. To be able to differentiate between different species, one would need to examine the signal properties more closely. This is demonstrated in [12], where several features are utilised, including standard deviation, entropy, spectrum centroid, and locations and amplitudes of peaks in the frequency domain.

In previous work, a proof of concept was shown, where an elephant was detected and tracked using a geophone network [9]. The major flaw in their work is the detection algorithm, as it was not validated against animals coexisting with elephants, raising concerns about false alarms. In this paper, we address these shortcomings by refining the detection algorithm, ensuring it accurately discriminates elephant footsteps from other species' footsteps in their habitat. This paper is based on the master's thesis [13], which is a continuation of the work in [9].

The paper is outlined as follows. In Section II, the proposed event detection and elephant footstep detection algorithms are presented. The DOA estimation method is outlined in Section III, while Section IV presents the field trials and the experimental setup. Section V describes the elephant footstep analysis, identifying the detection parameters, the detection performance as well as the DOA results. Finally, the conclusions are presented in Section VI.

II. Elephant Footstep Detection

The footstep of an elephant has unique characteristics which can be used to differentiate elephant footsteps from other species' footsteps. This section describes the methodology for the detection algorithm, which is based on unique signal features that the elephant's footstep exhibits.

A. Signal model

It is assumed that geophone i measures the signal

$$y_i(t) = \sum_n z_i^n(t) + e_i(t), \tag{1}$$

where $z_i^n(t)$ is the nth event for geophone i, which could be an elephant footstep, and $e_i(t)$ is ambient noise. An illustration of the signal model with two events is shown in Fig. 2. For simplicity, the noise is neglected in the figure.

B. Footstep extraction

An important aspect of the detection algorithm is to analyse the characteristic of an elephant's footstep. As our method requires an extracted segment of the geophone signal, it is important that the segments are of appropriate size; large enough to contain an elephant's footstep, and small enough to save processing power.

An elephant's footstep has been shown to last around 100–250 ms [14]. To include the whole footstep, while minimising the processing power, the segment length is chosen to be 270 ms. The segments are implemented as a sliding window to cover the elephant's footstep, and if a step is detected, the next segment is chosen to begin where the previous one ended. This is done to prevent having data from an old step and thus causing detection of the same step twice.

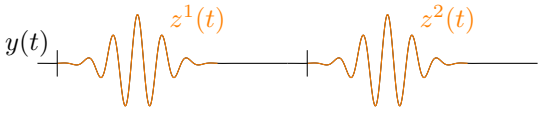


Fig. 2: An overview of the signal model depicting two distinctive events highlighted in orange. The first event is denoted as $z^1(t)$, while the second event is represented as $z^2(t)$.

C. Event detection

To determine if a segment includes an elephant's footstep, the algorithm first detects if there is an event, which could be any type of signal that is not ambient noise. This is done by calculating the kurtosis of the segment, defined as the fourth standardised moment

$$Q[k] = \frac{\sum_{l=k-N+1}^{k} (\bar{y}[l] - \mu[k])^4 / N}{\left(\sum_{l=k-N+1}^{k} (\bar{y}[l] - \mu[k])^2 / N\right)^2}, \tag{2}$$

where $Q[k]$ is the kurtosis at sample k, N is the number of samples in the window, μ is the calculated mean

$$\mu[k] = \sum_{l=k-N+1}^{k} \bar{y}[l], \tag{3}$$

and $\bar{y}[l]$ is the summed amplitude of all geophones at sample l

$$\bar{y}[l] = \sum_{i=1}^{M} y_i[l], \tag{4}$$

where M is the number of geophones. Kurtosis is a measure of the "tailedness" of a distribution, which is how often outliers occur [15]. This means that if a signal differs from Gaussian noise, it will have a steeper peak around the mean. Therefore, kurtosis is a good statistic measure to differentiate an event from ambient noise. Note that Gaussian noise has a kurtosis of 3.

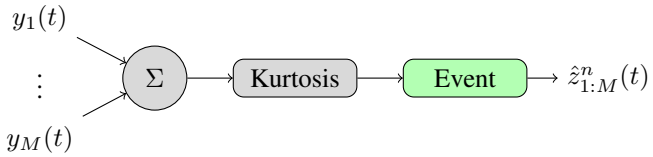
A threshold is used for the extraction of events. If the threshold is met, the segment is chosen to begin δ seconds before the threshold was hit, in order to contain the tail of the footstep. The extracted nth event for geophone i can therefore be denoted as

$$\hat{z}_i^n = y_i[k_n - \delta f_s : k_n + N - \delta f_s], \quad Q[k_n] \geq T, \tag{5}$$

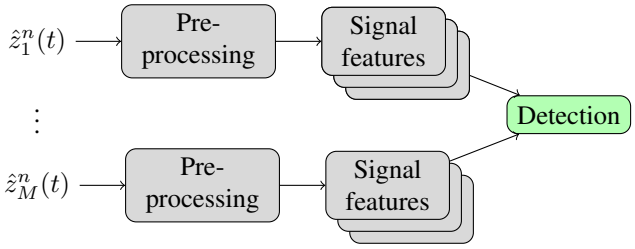
where T is the threshold and f_s is the sampling frequency. An overview of the event detection is shown in Fig. 3(a).

D. Preprocessing

The extracted event is thereafter preprocessed based on the characteristic of an elephant footstep. An elephant's footstep has a frequency range that rarely exceeds 30 Hz, with a mean frequency of 20 Hz [14]. Also, it has been noted that a noise component sometimes appears around 50 Hz [9]. To isolate the



(a) An overview of the different modules of the event detection.



(b) An overview of the different modules of the footstep detection.

Fig. 3: A schematic overview of the different stages of the detection algorithm.

interesting part of the signal, and reduce the influence of noise when analysing the characteristics of the signal, a band-pass Butterworth filter, with cutoff frequencies 4 and 45 Hz, and order 6 is applied to the signal. The reason for a band-pass filter, rather than a low-pass filter, is to reduce low frequency noise and sensor biases.

Before extracting the characteristics of the signal, the signal is also normalised, such that the power of the segment is 1. This approach serves to mitigate the potential impact of variations in range, size, and gender among the elephants.

E. Signal characteristic

To determine if the detected event corresponds to an elephant footstep, distinctive characteristics of the elephant footstep are utilised. These characteristics include the standard deviation, the frequency peak, the spectral centroid, and the frequency distribution of the signal. Subsequently, these attributes are compared to those typical of an elephant footstep to determine the presence of such a signal.

An overview of the detection algorithm's various components is provided in Fig. 3(b). While challenging to quantify precisely, the order of the features has been selected such that the feature with the highest capacity to eliminate false detections is evaluated first. This strategy is implemented within a decision tree framework to streamline computations and avoid unnecessary calculations.

For a signal to be classified as a footstep, it must successfully comply with all the features. Additionally, in order for a detection to be registered, the criteria for signal features must be satisfied across all geophones.

1) Standard deviation: The first feature that is analysed is the standard deviation of the preprocessed signal in the time domain

$$\sigma_i = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\hat{z}_i[k]-\mu)^2}. \tag{6}$$

The standard deviation varies among signals in the time domain and serves to distinguish between them. If the standard deviation falls within predefined thresholds, it passes the test; otherwise, it is discarded.

2) Frequency peak: The next characteristic of the signal to be examined is its major frequency peak. To calculate the frequency content, a *discrete Fourier transform* (DFT) is applied on the preprocessed signal, wherein a Hamming window is applied to prevent leakage. Subsequently, the frequency with the highest magnitude is determined. If the main frequency peak falls within specified threshold values, it passes the test; otherwise, it is discarded.

3) Spectral centroid: The spectral centroid, derived from the same DFT as the frequency peak, determines the frequency around which the signal's spectrum is centered. It is computed as a weighted mean of the frequency content of the signal

$$C_i = \frac{\sum_{n=1}^{\frac{N}{2}+1} f[n]|Z_i[n]|}{\sum_{n=1}^{\frac{N}{2}+1}|Z_i[n]|}, \tag{7}$$

where C_i represents the centroid of geophone i, N denotes the number of samples, $Z_i[n]$ is the amplitude corresponding to bin n and geophone i. Lastly, $f[n] = {}^{n f_s}/_{N}$ indicates the frequency corresponding to bin n [16]. If the centroid falls within predetermined thresholds, it passes the test; otherwise, it is discarded.

4) Frequency distribution: Lastly, the energy of the high frequencies is compared to that of low frequencies. This is accomplished by applying a 6th order Butterworth high-pass filter with a cutoff frequency of 55 Hz to the raw signal segment. The energy of the low frequencies, extracted from the preprocessed signal described in Section II-D, is then divided by the energy of the high frequencies to obtain a ratio

$$\text{Ratio}_i = \sqrt{\frac{\sum_{k=1}^{N}\left|\left(\mathcal{H}^{\text{low}} \star \hat{z}_i\right)[k]\right|^2}{\sum_{k=1}^{N}\left|\left(\mathcal{H}^{\text{high}} \star \hat{z}_i\right)[k]\right|^2}}, \tag{8}$$

where $\star$ represents the convolution operation, and $\mathcal{H}^{\text{low}}$ and $\mathcal{H}^{\text{high}}$ denote the low-pass and high-pass filters respectively. If the ratio exceeds a specified threshold, the test is passed. While a low ratio may result from high-frequency noise potentially interfering with the detection algorithm, the energy of footsteps typically outweighs that of noise.

Assessing the ratio between low and high frequencies serves two purposes. Firstly, it efficiently filters out numerous animals besides elephants. Secondly, if an elephant is detected, but the segmentation is inadequate, resulting in a segment primarily comprising ambient noise with only a small part of the elephant footstep, the segment typically exhibits higher energy at higher frequencies than usual. Discarding such segments early in the process is beneficial for subsequent direction estimation tasks.

III. DOA Estimation

By utilising the detected signal from multiple geophones, it becomes feasible to accurately determine the direction of an

elephant's footsteps relative the array. This section presents the fundamental principles for estimating the DOA from multiple signals, highly inspired by our previous work [9].

A. Signal model

It is assumed that the detected signal from Section II for geophone i can be expressed as

$$z_i[k] = s\left[k - \tau_i\left(\phi[k]\right)\right] + e_i[k], \tag{9}$$

where

- $s[k]$ is the seismic signature of an elephant footstep.
- $\phi[k]$ is the DOA at sample k. The array is assumed to satisfy a far-field assumption, *i.e.*, the DOA is the same for all geophones.
- $\tau_i(\phi)$ is the geometric discrete time-delay to geophone i with respect to the local origin of the array.
- $e_i[k]$ are all non-interesting signals measured by geophone i, *e.g.*, electrical noise, sensor bias or other ambient signals.

B. Delay and sum

In order to estimate the DOA of the elephant's footstep, the delay-and-sum beamforming technique is applied to $z_i[k]$ [17]. This method utilises the geometry of the array and computes the DOA by evaluating the power of the combined signals across all potential arrival angles. The maximum of the power corresponds to the estimated angle of arrival

$$z_{DS}^{\phi}[k] = \frac{1}{M}\sum_{i=1}^{M} z_i[k + \tau_i(\phi)], \tag{10a}$$

$$V_N(\phi) = \sum_{k=1}^{N} \left|z_{DS}^{\phi}[k]\right|^2, \tag{10b}$$

$$\hat{\phi} = \arg\max_{\phi} V_N(\phi). \tag{10c}$$

To compute the variance and refine the DOA estimate, a second degree polynomial is fit locally around the maximum

$$V_N(\phi) \approx a + b\left(\phi - \hat{\phi}\right) + c\left(\phi - \hat{\phi}\right)^2, \tag{11}$$

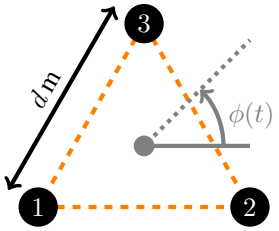
where the parameters a, b and c are estimated using the *least squares* (LS) method. The range of angles ϕ to be included around the maximum $\hat{\phi}$ is crucial to get a good estimate of the variance. Once this polynomial is estimated, it can be shown that a refined estimate of the angle $\hat{\hat{\phi}}$ and its variance can be computed as

$$\hat{\hat{\phi}} = \hat{\phi} - \frac{b}{2c}, \tag{12a}$$

$$\mathrm{Var}(\hat{\hat{\phi}}) = \frac{\frac{b^2}{4c} - a}{Nc}, \tag{12b}$$

where $\hat{\hat{\phi}}$ is the refined DOA estimate and N is the number of samples in the signal.

(a) Hardware overview of the complete system with the following items: **1)** three geophones, **2)** a box that contains the microcontroller, ADC, battery, GNSS and LTE antenna, and **3)** a solar panel.



(b) The setup of the geophone network. The distance between the geophones is d meters and the DOA angle is denoted $\phi(t)$.

Fig. 4: The hardware components used and the setup of the geophone network.

IV. Experimental Setup

This section outlines the hardware design and details the experimental setup employed, which is similar to that used in our previous work [9].

A. Hardware

To evaluate the algorithms, a sensor prototype was devised, comprising three SM-24 Geophone Elements, each with a bandwidth ranging from 10 to 240 Hz [18]. The analog signals from the geophones are passed through an *analog to digital converter* (ADC), ADS1256 [19], before being processed by a microcontroller board, LILYGO® TTGO T-SIM7000G ESP32 board [20]. The ADC includes a programmable gain amplifier (PGA) with adjustable values ranging from 1 to 64, and for this work, the PGA was set to 32. The entire system is powered by a combination of solar panels and batteries. An illustration of the hardware components is provided in Fig. 4(a).

B. Geometric Delay

The geophones were arranged in an equilateral triangle configuration with a side length of 4 m, forming a *Uniform Circular Array* (UCA) as depicted in Fig. 4(b). This configuration yields geometric discrete time-delays with respect to the array's center given by

$$\tau_i(\phi) = -\frac{df_s}{v\sqrt{3}}\cos(\phi + \psi_i), \quad i = 1, 2, 3, \tag{13}$$

where d represents the distance between geophones, v is the propagation speed of the elephant's footstep, and ψ_i denotes the angle to the ith geophone relative to the array's center.

C. Sampling

The geophone operates as an electromagnet, generating a voltage when displaced. To utilise this voltage, it has to be converted into a digital signal by the ADC. The ADC theoretically offers a sampling frequency range of 2.5 Hz to 30 kHz, offering a balance between resolution and sampling speed. However, considering that the upper bandwidth of the geophones is 240 Hz, a sampling frequency exceeding 480 Hz is unnecessary. Additionally, since Rayleigh waves seldom exceed 200 Hz, a minimum sampling frequency of 400 Hz is

Fig. 5: The setup of the geophone network at Kolmården. The DOA angle is denoted $\phi(t)$. The elephant walks in front of the geophone network at a distance of 8–30 m.

necessary to prevent aliasing. Due to hardware constraints, the actual sampling frequency ultimately reached 474 Hz.

D. Experiments

Data collection was primarily conducted at Kolmården Wildlife Park using the developed prototype. Initially, the wave propagation speed was measured by performing foot stomps along the baseline for one pair of geophones. Cross-correlation was employed to retrieve the time delays from 10 different measurements, resulting in a mean wave propagation speed of $v = 161.7$ m/s with a standard deviation of 12.3 m/s. It is worth noting that the wave propagation speed can vary significantly depending on the ground composition, thus a recalibration would be necessary for each new environment.

During the recording of the data set, a male Asian elephant walked from one side of the enclosure to the other, corresponding to an angle ϕ ranging from 20 to 140 degrees. The elephant maintained a distance of approximately 8 to 30 meters from the array, as illustrated in Fig. 5.

The accuracy of DOA estimates were validated using a 360-degree camera placed in the center of the geophone array. To maintain fidelity to real angles, the footage was projected using an equirectangular projection. Additionally, the three geophones served as reference angles in the video footage. The video and geophone data were synchronised by executing a foot stomp, observable in both the video footage and the geophone data.

Footstep measurements of various animals were conducted both at Kolmården Wildlife Park and in Kenya. Footsteps were counted from the video footage, and data from both locations are outlined below.

1) Kolmården Wildlife Park: Data was collected from several species at Kolmården Wildlife Park, with a primary focus on a male Asian elephant (*Elephas Maximus*). Additionally, data was collected from a rhinoceros, kulans, and Asian camels.

2) Kenya: Datasets were gathered from two locations within Tsavo West National Park, Kenya: waterhole four at Ngulia Rhino Sanctuary and a waterhole outside Kilaguni Serena Safari Lodge. Elephant and human footsteps were recorded at Ngulia, while water buffalo footsteps were collected at Kilaguni.

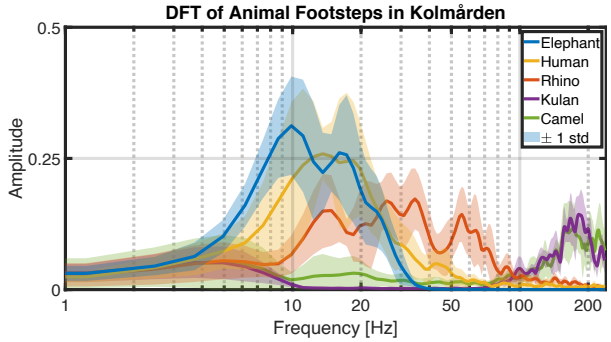


Fig. 6: A DFT comparison between elephant, human, rhino, kulan and camel footsteps at Kolmården Wildlife Park. The solid lines represent a mean value of ten footsteps from the animal. The shaded areas indicate one standard deviation from the mean.

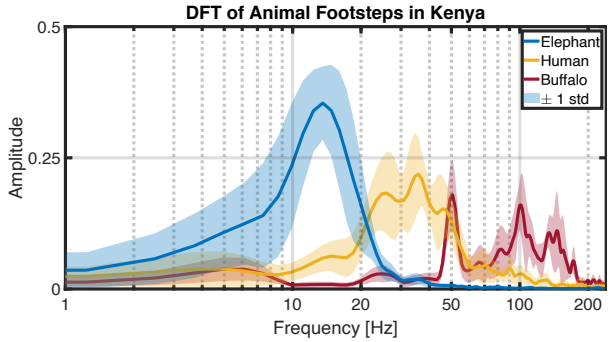


Fig. 7: A DFT comparison between elephant, human, and water buffalo footsteps in Kenya. The solid lines represent a mean value of ten footsteps from the animal. The shaded areas indicate one standard deviation from the mean.

V. Results

In this section, results regarding the detection method and DOA estimation are presented.

A. Animal footstep analysis

The extracted segments from each animal's footsteps were analysed in both the time and frequency domains. Fig. 6 and Fig. 7 shows a comparison of the frequency content between different animal footsteps in Kolmården and Kenya, respectively. The figures illustrate the mean normalised DFT of ten footsteps per animal, showcasing distinctive frequency characteristics. Each signal segment is 128 samples long and is zero-padded with 256 zeros to increase the resolution of the DFT.

B. Choosing detection parameters

Given the abundance of data from Kolmården compared to Kenya, the detection algorithm was tailored to detect elephants primarily at Kolmården. Initially, the parameter δ and the kurtosis threshold were tuned to capture an entire footstep as well as to distinguish between events and ambient noise. By calculating the kurtosis of ambient noise and for an elephant footstep, as seen in Fig. 8, the threshold was set to 4.3. This threshold is low enough to react to the elephant's footstep, and high enough to be unresponsive to ambient noise. To capture

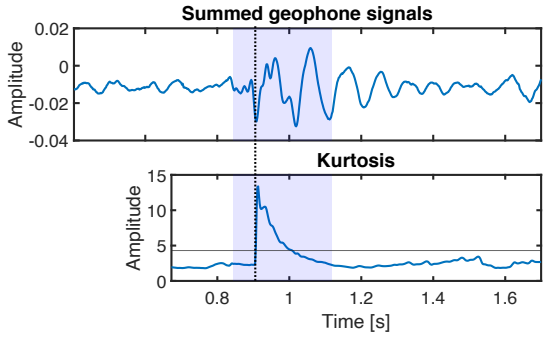


Fig. 8: The summed signals from an elephant's footstep, and the corresponding kurtosis. The threshold is illustrated with the thin solid line, the dotted line indicates where the threshold is reached and the shaded area denotes the extracted footstep.

TABLE I
SIGNAL FEATURE VALUES, EXTRACTED FROM 109 ELEPHANT FOOTSTEPS AT KOLMÅRDEN WILDLIFE PARK AND 10 ELEPHANT FOOTSTEPS AT NGULIA RHINO SANCTUARY IN KENYA.

	Kolmården		**Ngulia**	
Feature	Mean	σ	Mean	σ
Standard deviation	0.08821	6.5e-04	0.08812	9.3e-04
Frequency peak (Hz)	13.2	3.72	12.9	3.15
Spectral centroid (Hz)	19.1	3.62	18.3	3.53
Frequency ratio	603.1	478.9	48.1	27.4

the entire elephant footstep, including the tail in the beginning, δ was set to 60 ms, corresponding to 28 samples.

The signal feature parameters are calculated from 109 footsteps, collected from a single male elephant at Kolmården Wildlife Park. The resulting signal features, as described in Section II, are presented in Table I.

Using the metrics from Table I, thresholds can be determined. Three different sets of thresholds are evaluated for the three first features. The thresholds have been chosen as multiples of the standard deviation (1, 1.5 and 2), of each respective feature, above and below the mean of the parameters. Additionally, the threshold for the frequency ratio is set to 15 at all times. This is low enough to allow clear elephant step segments to pass, while discarding animals with high frequency steps. It is also high enough to not be massively effected by high frequency noise.

C. Detection performance

To evaluate the performance of the detection algorithm, video footage was used to count visible footsteps. The performance is only tested on the elephants at Kolmården, and not Kenya. This is because the elephant data from Kenya already passed through a lighter version of the detection method on the hardware. Therefore, assessing the performance of the detection algorithms on this data would not be an accurate representation of the actual performance. The performance is assessed with respect to

- *True Positives* (TP) — Detection when an elephant footstep has occurred.
- *True Negatives* (TN) — No detection when anything other than an elephant footstep has occurred.

TABLE II
DETECTION RESULTS FROM DIFFERENT ANIMAL FOOTSTEPS COLLECTED AT KOLMÅRDEN WILDLIFE PARK. THE FIRST VALUE IN EACH CELL REPRESENTS 1 STANDARD DEVIATION, THE MIDDLE 1.5 STANDARD DEVIATIONS AND THE LAST 2 STANDARD DEVIATIONS (1σ / 1.5σ / 2σ).

Animal	**TP**	**TN**	**FP**	**FN**
Elephant	17 / 31 / 35	0 / 0 / 0	0 / 1 / 1	40 / 26 / 22
Human	0 / 0 / 0	57 / 57 / 54	0 / 0 / 3	0 / 0 / 0
Rhino	0 / 0 / 0	57 / 54 / 53	0 / 3 / 4	0 / 0 / 0
Camel	0 / 0 / 0	57 / 57 / 57	0 / 0 / 0	0 / 0 / 0
Kulan	0 / 0 / 0	49 / 49 / 49	0 / 0 / 0	0 / 0 / 0
Total	17 / 31 / **35**	**220** / 217 / 213	**0** / 4 / 8	40 / 26 / **22**

- *False Positives* (FP) — False alarm, detection when elephant footstep has not occurred.
- *False Negatives* (FN) — Missed detection when an elephant footstep has occurred.

The models are also evaluated using accuracy, recall, precision and F1 score

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{14a}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{14b}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{14c}$$

$$\text{F1 score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{14d}$$

F1 score is a harmonic mean of the recall and precision of a model. This gives a good overview of the performance of the algorithm. All metrics ranges from 0 to 1, where a higher value is better [21].

The models are evaluated on datasets different from the ones that contain the training data. The elephant validation data, recorded in 2022, is taken from a single male elephant, walking about 15–40 meters from the array [9]. During the data collection for the rhinos one of the geophones was malfunctioning, and hence the detection criteria are only applied to the two working channels. The footsteps were counted from video footage, to make the amount of missed detections evident. All models have been validated on the same datasets. Each animal validation is done with one or more datasets where the geophones have recorded continuous seismic data, containing 57 footsteps of each respective animal, except for the kulan, where only 49 footsteps were recorded. Table II shows the detection results for the different animals footsteps. Notably, the false positive from the elephant data is from in between the footsteps. That is, it originates from the rumbles of an already detected footstep, resulting in the algorithm detecting the same footstep twice. The accuracy, recall, precision and F1 score of the models are presented in Table III.

D. Detection discussion

Overall, the detection method works well. It is evident from Fig. 6 and Fig. 7 that some animals are easier to differentiate than others. In Fig. 6, the camel and kulan DFTs have frequencies that are significantly higher, compared to the

TABLE III
EVALUATION SCORES CALCULATED FROM THE DETECTION RESULTS IN TABLE II.

Standard deviation(s)	Accuracy	Recall	Precision	F1 score
1	0.86	0.30	**1**	0.46
1.5	**0.89**	0.54	0.89	0.67
2	**0.89**	**0.61**	0.81	**0.7**

other three animals. This directly translates into the results of Table II, where no kulan nor camel footsteps have been detected for any of the thresholds. An unexpected result is that more rhino footsteps were detected than human footsteps, which is not obvious from Fig. 6. One explanation for this could be the fact that one of the geophones was not working during these measurements, making false positives more likely.

Looking at Table III, one could conclude that the algorithm works best with the thresholds at two standard deviations. It has the highest accuracy, recall and F1 score. However, one could argue that for the elephant detector, it is important to have high precision. If the detectors are deployed in the field, and sending reports to park rangers, frequent false alarms from human footsteps could cause the rangers to stop listening to it, rendering the device useless. Therefore, using two standard deviations as thresholds is too high for the intended purpose. The same could be said about setting the thresholds at 1.5 standard deviations, as it also had some false alarms from rhino footsteps. However, knowing that the rhino data were obtained with only two out of three geophones makes these false positives less significant. The rarity of rhinos in the wild also makes it very unlikely to detect one, and if one is detected, this information would also be of importance to the park rangers. Therefore, 1.5 standard deviations seem to be a good trade-off between recall and precision, giving it a high F1 score, as well as having high accuracy.

Although there is not an abundance of data from Kenya, by comparing the data with the one from Kolmården in Table I, one could draw the conclusion that the detection algorithm would work well for African elephants as well. The metrics are very similar, with the most deviating one being the frequency ratio. It is however difficult to come to a certain conclusion on this, due to the lack of data.

Comparisons with existing algorithms showcased notable improvements, particularly in reducing false positives from other animal footsteps, as seen in Table IV. Based on this small sample of 10 footsteps of each animal, it is shown that the proposed algorithm produces no false positive from another animal, while the algorithm in [9] generates a false alarm for both a human and a rhino footstep. As discussed, the footstep from the rhino would also be relevant to detect, while the human footstep is more bothersome, as humans could potentially walk close by the sensor array in the intended application. The evaluation score for the two methods are shown in Table V.

TABLE IV
DETECTION RESULTS FROM 10 DIFFERENT ANIMAL FOOTSTEPS COLLECTED AT KOLMÅRDEN WILDLIFE PARK USING THE PROPOSED ALGORITHM AS WELL AS THE METHOD FROM [9].

	Proposed algorithm				**Algorithm from [9]**			
Animal	TP	TN	FP	FN	TP	TN	FP	FN
Elephant	8	0	1	2	10	0	8	0
Human	0	10	0	0	0	9	1	0
Rhino	0	10	0	0	0	9	1	0
Camel	0	10	0	0	0	10	0	0
Kulan	0	10	0	0	0	10	0	0
Total	8	**40**	**1**	2	**10**	38	10	**0**

TABLE V
EVALUATION SCORES CALCULATED FROM THE DETECTION RESULTS IN TABLE IV OF THE PROPOSED ALGORITHM AND THE METHOD IN [9].

Metod	Accuracy	Recall	Precision	F1 score
Proposed	**0.94**	0.8	**0.89**	**0.84**
From [9]	0.83	**1**	0.5	0.67

E. DOA measurements

Utilising the identified segments from the detection algorithm of the three geophones, it is possible to estimate the angle to the elephant. In Fig. 9, the DOA estimates from a walking elephant at Kolmården is illustrated. The estimates $\hat{\phi}$ have been refined with the method described in Section III-B, with an interval of 20 degrees around the maximum $\hat{\phi}$. The elephant walked back and forth in front of the geophone network with a distance varying from 8 to 30 meters. The measurements are validated for all the elephant steps except the steps between approximately 10 and 30 degrees where a bush was located. It is seen that the refined DOA estimates follows the blue line of the elephant to a great extent. There are also a few outliers, 9 out of 112 estimates, outside the figure interval, *i.e.*, $\hat{\phi} < 0°$ or $\hat{\phi} > 180°$.

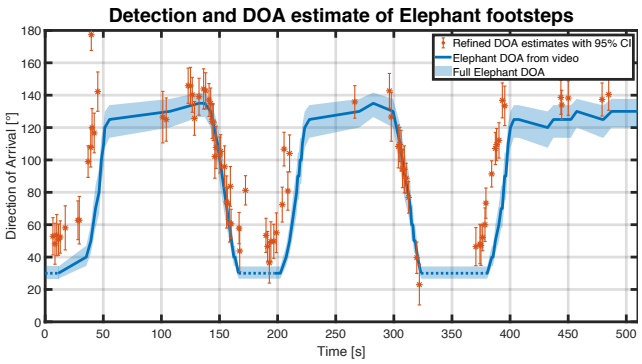


Fig. 9: Refined DOA estimates of a walking elephant at Kolmården, using the segments received from the detection algorithm with 1.5 standard deviations. The solid blue line represents DOA estimates from the 360 camera, the dotted blue line indicates when the elephant walked behind the bush, and the red stars is the refined DOA estimates, $\hat{\phi}$, with 95% confidence interval. The shaded blue area is the approximate DOA of the full elephant, including front foot and hind foot.

VI. Conclusion

This article introduces a novel method for detecting elephant footsteps through a geophone network. Additionally, it accurately estimates the direction of the elephant using a delay-and-sum beamformer.

The detection method achieves an accuracy of 89 % and successfully detects 54 % of the elephant footsteps. Moreover, it demonstrates an ability to differentiate elephants from other species while maintaining a low false alarm rate.

The system reliably estimates the direction of an elephant relative to a geophone array, with the majority of the estimates closely aligned with the elephant's position. Incorporating a Kalman filter could further enhance tracking accuracy by effectively managing outliers.

Several avenues for future improvement exist. Enhancements to the detection algorithm, such as incorporating additional signal features or exploring machine learning approaches, could improve accuracy. However, the pursuit of machine learning methods necessitates extensive data sets, which is currently limited. Furthermore, extending the system to detect and classify other animals would enhance its versatility.

Another promising direction involves deploying an additional network of geophones in proximity to enable not just the estimation of elephant direction but also its precise position. This extension could significantly enhance the system's capabilities and applications.

References

[1] E. H. Enukwa, "Human-elephant conflict mitigation methods: A review of effectiveness and sustainability," *Journal of Wildlife and Biodiversity*, vol. 1, no. 2, p. 69–78, Oct. 2017.

[2] P. Stephenson, "The future for elephants in Africa," in *Terrestrial Ecoregions of Africa and Madagascar: A Conservation Assessment*, N. Burgess *et al.*, Eds. Washington DC, USA: Island Press, 2004, pp. 133–136.

[3] L. J. Shaffer, K. K. Khadka, J. Van Den Hoek, and K. J. Naithani, "Human-elephant conflict: A review of current management strategies and future directions," *Frontiers in Ecology and Evolution*, vol. 6, 2019.

[4] J. D. Wood, C. E. O'Connel-Rodwell, and S. L. Klemperer, "Methodological insights: Using seismic sensors to detect elephants and other large mammals: a potential census technique," *Journal of Applied Ecology*, vol. 42, no. 3, pp. 587–594, 2005.

[5] J. L. Wijayaraja, J. L. Wijekoon, M. Wijesundara, and L. J. M. Wickramasinghe, "Towards Long Range Detection of Elephants Using Seismic Signals; A Geophone-Sensor Interface for Embedded Systems," *IEEE Access*, 2024.

[6] C. G. Sammis and T. L. Henyey, *Geophysics Field Measurements*. Academic Press, 1987.

[7] C. E. O'Connell-Rodwell, B. T. Arnason, and L. A. Hart, "Seismic properties of Asian elephant (Elephas maximus) vocalizations and locomotion," *The Journal of the Acoustical Society of America*, vol. 108, no. 6, pp. 3066–3072, 12 2000.

[8] L. Rayleigh, "On waves propagated along the plane surface of an elastic solid," *Proceedings of the London Mathematical Society*, vol. s1-17, no. 1, p. 4–11, Nov 1885.

[9] G. Zetterqvist, E. Wahledow, P. Sjövik, F. Gustafsson, and G. Hendeby, "Elephant DOA estimation using a geophone network," in *2023 26th International Conference on Information Fusion (FUSION)*, 2023, pp. 1–6.

[10] P. Anghelescu, G. V. Iana, and I. Tramandan, "Human footstep detection using seismic sensors," in *2015 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2015, pp. AE–1–AE–2.

[11] R. Samkaria, R. Singh, A. Gehlot, R. Pachauri, A. Kumar, N. K. Singh, and K. Rawat, "IOT and XBee triggered based adaptive intrusion detection using geophone and quick response by UAV," *International Journal of Engineering & Technology*, vol. 7, no. 2.6, pp. 12–18, 2018.

[12] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, and P. Zhang, "Indoor person identification through footstep induced structural vibration," in *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 81–86.

[13] D. Goderik and A. Westlund, "Detection and Tracking of Elephants using Seismic Direction of Arrival Estimates," 2023.

[14] C. O'Connell-Rodwell, B. Arnason, and L. Hart, "Seismic properties of Asian elephant (Elephas maximus) vocalizations and locomotion," *The Journal of the Acoustical Society of America*, vol. 108, no. 6, pp. 3066–3072, 2000.

[15] G. Koç and K. Yegin, "Footstep and vehicle detection using slow and quick adaptive thresholds algorithm," *International Journal of Distributed Sensor Networks*, vol. 9, no. 10, p. 783604, 2013.

[16] G. Peters, "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," ICARM: Paris, France, Tech. Rep., 2004.

[17] H. L. V. Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, Ltd, 2002.

[18] *SM-24 Geophone Element*, SENSOR Nederland b.v, 2006.

[19] *Very Low Noise, 24-Bit Analog-to-Digital Converter*, Texas Instruments, 2003, revised 2013.

[20] R. N. Tutorials, "Getting started with LILYGO T-SIM7000G ESP32 (LTE, GPRS, and GPS)."

[21] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, Aug 2015.