Elephant Censusing via Geophone Arrays: A Visual Approach for Linear Arrays

Gabriel J Chandler*, Ozgur Izmirli†, Caitlin O'Connell-Rodwell‡, and Jason Wood§
*Department of Mathematics, Connecticut College, New London, CT 06320-4196 USA, Email: gjcha2@conncoll.edu
†Department of Computer Science, Connecticut College, New London, CT 06320-4196 USA
‡Department of Otolaryngology, Stanford University, Stanford, CA 94305-5739 USA
§The Whale Museum, Friday Harbor, WA 98250-0945 USA

***Abstract*—The problem of censusing elephants via geo-phone recordings of their footsteps along a linear array positioned along a path leading to a watering hole is considered. We propose an intuitive and graphically motivated technique that delineates passings of herds of elephants. Additionally, a measure of dispersion of the localization estimates is proposed that correlates to the size of the passing herd. We illustrate the methods using both real data and a simulation study.**

I. INTRODUCTION

We consider in particular the problem of censusing elephant populations in forested regions of their habitat [Woods, et al. 2005]. The general problem of censusing includes problems of localization, tracking and segmentation of groups of gaited animals in their own habitat. In this work we focus on a particular case where groups of elephants move along a linear sensor array embedded in a noisy environment . The methodology presented below addresses the problems of localization, tracking and segmentation, and in addition gives indication to the size of the passing group. Other applications include tracking passage of herds to analyze behavioral patterns and time segmentation to extract sections of noisy data for further analysis such as footstep detection and gait modeling.

An accurate estimate of the size of an elephant population requires the ability to estimate the number of elephants passing in a group, as well as classification of passing animals as either elephants or otherwise. In this paper, we concentrate on the methods of segmentation of long term data into sections which consist of information pertaining to either a single animal or a group. With proper segmentation, problems such as counting footsteps and classification of animal type (see [1]) is possible, though beyond the scope of the present work. We premise that any method used for solving such problems should be done using the entire array data. As such, we look to segment the data into sections for which the entirety of the array is being dominated by a single group of elephants. We use seismic detection for minimal human intervention to the natural environment.

There has been significant research in the field of localization with a wide variety of applications ranging from acoustic to radar problems. The nature of the sources and properties of the environment dictate different approaches to the detection of the required information [2]-[4]. In source localization terminology, our problem can be viewed as localization of a group of narrowband, near-field, moving seismic sources with unknown signal propagation speed in a reverberant environment. In addition, the sources are impulsive and not necessarily strictly periodic. In our case, the main difficulty in localization arises from seismic properties such as unknown reverberation patterns and rapid energy dissipation.

We refer to a small sampling of work related to our method in this paper. Knapp and Carter proposed a generalized correlation method for two-channel time delay estimation [5]. The method used a maximum likelihood estimator and a prefilter to suppress noise power. Mainwaring et al. used a tiered sensor network architecture for habitat monitoring of seabird nesting behavior [6]. Gannot and Dvorkind proposed a two-level method for acoustic source localization with the first level estimating pairwise time delays and the second level combining this information for a microphone array [7]. Other research on acoustic source localization concentrates on arrangement of microphone arrays [8], [9] and tracking multiple speakers [10].

Section 2 contains a description of the data and preprocessing, Section 3 describes the procedure for localization, section 4 discusses the parameterization used in estimating location, and results related to estimating group size are given in section 5. A conclusion of the methods and results is found in section 6.

II. PREPROCESSING THE ARRAY DATA

The input consists of 13 channels of sampled data obtained from equidistant geophones. Due to the fact that the data was extremely noisy we applied a spectral filter that would maximize the ratio of the energy output for footstep signals to all other spectrally unrelated ones. Each channel is filtered in the spectral domain with a filter that has a bandwidth spanning a typical footstep spectrum. This is implemented by calculating the sliding short-time Fast Fourier Transform of each channel. Each window that is 0.25 seconds long is multiplied with a Hann window prior to calculating the transform and successive windows are overlapped 87.5% in time. The resulting banded amplitude spectrum is then squared and summed over all bins to obtain the filtered signal energy for that window. The large dynamic range of the energy signal led us to consider compression in order to reduce the dynamic range. This was more useful for us to visually observe the

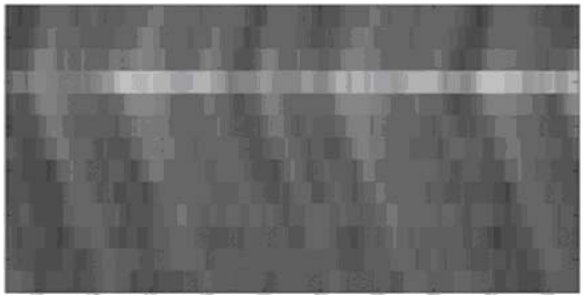

Fig. 1. Energy across channels.

interchannel relationships rather than having a computational benefit. Compression is done by raising the energy signal to a power x, with $x < 1$. That is, to achieve less difference between sections that have higher energy where the elephant is close to a geophone and those that are very low in energy where the elephant is far away. The method presented in this paper relies on the information supplied by this front-end which constitutes solely of energies in the 13 channels.

III. LOCALIZATION

In general, the sensor array consists of k equidistant co-linear sensors (k=13, in our case), letting s_i, $i = 1, ..., k$, denote location of the sensors along the array axis. In this study we assume sources are always on one side of the array axis where the path used by elephants lies.

Exploiting the information regarding the lag domain, we can parameterize the distance of the source from the array and location by fitting a ’v’ to the time-varying spectrum of the data. A single angle, related to the time delay of arrival (TDOA), can be related to the distance of the source from the array. The point at which the lag offsets change in sign represents the projection of the source onto the array axis. The change point is represented as the vertex of a horizontal ’v’ and is hence referred to as the v-tip. In Fig. 1, a section of the data consisting of a small number of footsteps is shown. We can see the ’v’s with the v-tip clearly near the 4th sensor. The ’v’ shape is an approximation to the theoretical lag pattern based on the assumption that the sources are relatively near to the array axis. Our experiments have shown that this assumption holds for our data but, in general, this lag pattern can be further parameterized if necessary. Assuming the source is moving parallel to the array, the angle will be constant and thus only a single value need be estimated in a chosen section of the input data. In the case the source is not moving parallel to the array, the angle can be treated as a function in time, and can be estimated, for example, by looking at possibly smaller sections of the data in time.

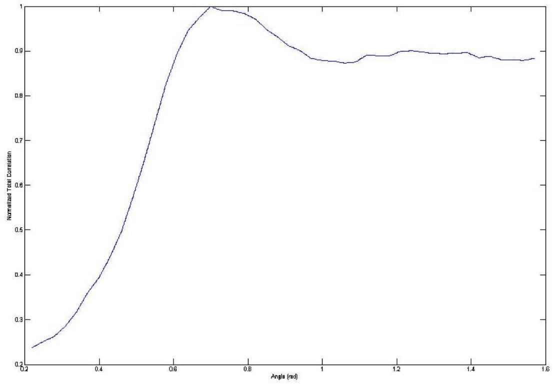

Fig. 2. $\sum_i \sum_j \hat{\rho}_{ij}$ as a function of angle θ.

First we estimate the optimal angle for each section. A given angle and v-tip determine the time lags at which to match energies for all channels. We consider the matrix of correlations between pairs of windows at these time lags across various angles. The length of the window is chosen such that it contains at least part of a footstep. This produces a series of correlation matrices of size k by k, where the i, j^{th} element is the correlation between the data segments of the i^{th} and j^{th} sensor. We define $\hat{\rho}_{ij}(v, \theta, s)$ as the correlation of the windowed data along the line determined by the v-tip location v and angle θ and time point s for the i^{th} and j^{th} sensors.

The estimated angle $\hat{\theta}$ maximizes the total sum of the correlation matrices across all time points in a section. The v-tip at each time point is then estimated using this angle as the position which has the largest sum of the correlation matrix at $\hat{\theta}$. Arbitrarily, we set the v-tips to be centered at the first channel. At the correct angle, the submatrix corresponding to the sensors beyond the true v-tip will tend to have large entries, in fact this submatrix will be the same as if we had known the correct v-tip location. At the incorrect angle, this will never happen, and any large entry in the correlation matrix will be due to chance. Thus, we expect that the selected angle will be close to the true angle, without specifying the v-tips. The estimate of the angle is given by:

$$\hat{\theta}(s) = \arg\max_{\theta} \sum_{t=s-m}^{t=s+m} \sum_{i=1}^{k} \sum_{j=1}^{k} \hat{\rho}_{ij}(1, \theta, t)$$

where $2m$ is the length of the section being analyzed centered at time s. For our particular situation it is safe to assume $\hat{\theta}(s)$ is not a function of s due to the orientation of the array along a path ensuring roughly parallel movement.

The peak in Fig. 2 shows the optimal angle for a single elephant passing. A clear maximum is seen despite not knowing the optimal v-tip locations. In this example, the section being analyzed contains the entire passage of a single elephant. Updating this calculation using estimates of the v-tip location, discussed below, could be done iteratively, allowing even more resolution of the optimal angle. Alternatively, the maximum energy as recorded by the sensors can be interpolated to arrive at an estimate of the v-tip position when the projected source location on the array axis is in between the two sensors on either end of the array. Fig. 3 shows the total correlation as a

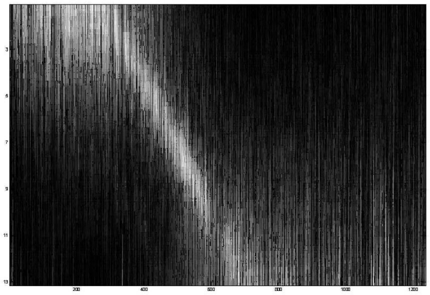

Fig. 3. Total correlation of energy along vs across position.

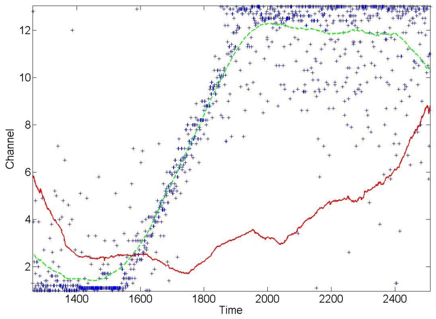

Fig. 4. Single elephant. Maxima of total correlation(points), moving average of the maxima and weighted measure of dispersion.

function of time and position:

$$g(s,v)=\sum_{i=1}^{k}\sum_{j=1}^{k}\hat{\rho}_{ij}(v,\hat{\theta}(s),s)$$

The lighter shaded regions correspond to higher total correlation. A line which corresponds to the passage of a single elephant is visible. Next, we find estimates for the v-tip locations by selecting the location that maximizes the total correlations:

$$\hat{v}(s)=\arg\max_{v\in[1,13]} g(s,v)$$

Fig. 4 shows $\hat{v}$, together with the moving average and the weighted measure of dispersion that is defined below.

IV. PARAMETRIZATION OF LOCATION BY ANGLE AND V-TIP

A. *Tracking*

The angle is related to the distance of the source from the array. For instance a source at infinity would result in the signal reaching all sensors simultaneously, and the resulting 'v' would have an angle of 90 degrees. The lower bound of the angle is determined by the propagation speed and the distance between sensors in the array. This is achieved when the source is co-linear with the array.

The location of the v-tips determines the location of the projection of the source on the array axis. Combined with angle, this determines the position of the source up to two possibilities, the side of the array the source is on is not identifiable with a linear array. However, we only consider one side of the array because we know that there is a path that the elephants use to access water.

We can further parameterize the movement of the source by estimating the function describing the location of the v-tips in time. In the simplest case, the source is moving at a constant speed parallel to the array. Thus, a linear fit to the v-tip estimates would allow us to estimate the speed of the animal, related to the slope of the line.

A non-parametric estimate of the function relating the v-tip locations to time is recommended when the assumption of constant speed is not likely to hold. A weighted moving average of the locations of the v-tip estimates would be appropriate for estimating this function. The weights involved are related to our confidence in the particular estimate of the v-tip location. This is quantified by the sum of the correlation matrix for the estimate of the v-tip location at time s. The weighted moving average is given by the following:

$$\tilde{v}(s)=\frac{\sum_{t=s-n}^{s+n}\hat{v}(t)g(t,\hat{v}(t))}{\sum_{t=s-n}^{s+n}g(t,\hat{v}(t))}$$

where the moving average is calculated over a window of size $2n$ with $n << m$.

B. *Segmentation*

Based on the localization and tracking estimates given above, we wish to find segments of data that contain reliable information. As $\tilde{v}(s)$ does not contain a measure of reliability, we further suggest methods to quantify the amount of coherence in a particular time window. Related to the weighted moving average, we consider measuring the amount of dispersion in the v-tip location estimates, again weighting the estimates by their associated total correlation matrix:

$$d(s)=\frac{\sum_{t=s-l}^{s+l}\left(\hat{v}(t)-\tilde{v}(t)\right)^2 g(t,\hat{v}(t))}{\sum_{t=s-l}^{s+l}g(t,\hat{v}(t))}$$

where the measure of dispersion is calculated over a window of length $2l$.

The segments that contain usable data can be obtained through selecting a threshold and using the sections with a value less than that threshold. The value of the threshold is, in general, environment dependent, and can be chosen by looking at the amount of variability in the data when there is no crossing occurring. The threshold can then be set empirically such that the probability of false detection is small.

V. RESULTS

We show the results obtained from real data collected by J. Woods and colleagues at Namibia's Etosha National Park. Starting with a simple case, a single elephant, in Fig. 3 and 4 we show the results of tracking and segmentation. The dispersion curve is normalized to a maximum of 13 to be shown on the same plot. Also, the maximum of the dispersion measure is referenced to white noise for it to be comparable for

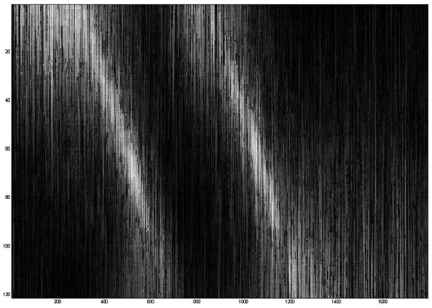

Fig. 5. Total correlation of energy along vs across position for simulated two elephant.

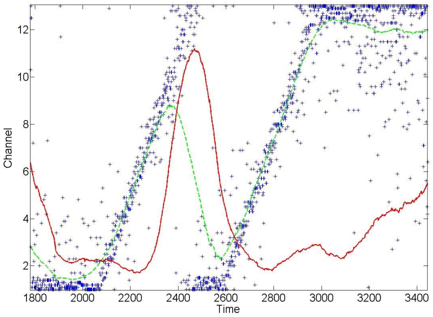

Fig. 6. Maxima of total correlation, moving average of the maxima and weighted measure of dispersion for simulated two elephant.

different data sets. The procedure for segmentation is simply to pick a threshold for the dispersion measure such that a workable segment is obtained. Lower thresholds will result in smaller segments of data, though with greater reliability. Next, we show how the dispersion measure could aid in segmenting individual passages of elephants. We do so by adding a shifted version of the single elephant case to itself and applying our method. In Fig. 5 we show the total correlation of the energy along vs across position and in Fig. 6 we show the weighted dispersion for the simulated two elephant case. Note the high value of dispersion while the maxima of total correlation switches from tracking one elephant to the other.

In Fig. 7 and 8 we show the same plots for a real case where two groups of elephants are passing. In fact, the figures show another group that is passing before the two groups come into focus. Finally, in Fig. 9 we show a single group of 12 elephants passing followed by another group.

A. Estimating Group Size

We point to the fact that the minima of the dispersion measure is correlated with the number of elephants in the passing group. The lower the minima, the lower the number of elephants in a group. For a single elephant, we can think of four sources moving along the array, each source being a foot. The variability in our measure is due to different sources producing the dominant signal in our window, as the sources are at slightly different locations. Thus, the estimated v-tip location will vary. For a larger group of elephants, there are both more sources and a greater spread amongst their locations. Consequently, we can expect to see greater variability in v-tip estimates over short time intervals.

To verify this relationship, we simulated groups of elephants ranging in size from 2 to 8 to understand the relationship between group size and the value of d at the time point at which the moving average crosses the middle sensor. We used the passage of a single elephant shown in Fig. 3 overlaid by random shifts with itself. Fig. 10 shows the values of d for the

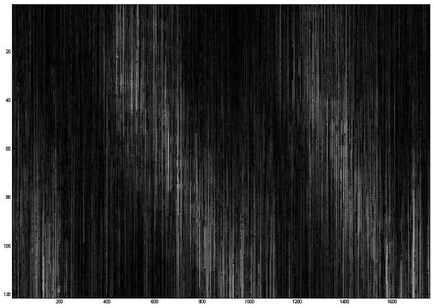

Fig. 7. Total correlation of energy along Vs across position for two elephant groups.

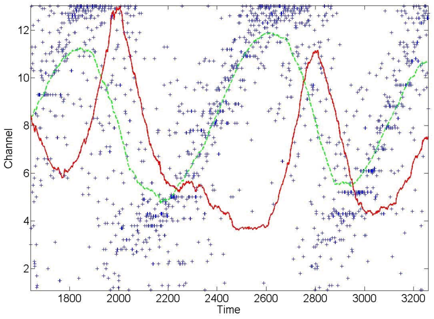

Fig. 8. Maxima of total correlation, moving average of the maxima and weighted measure of dispersion for two groups of elephants.

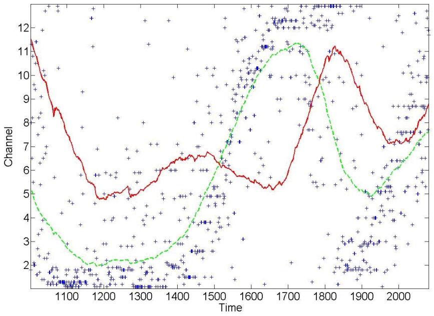

Fig. 9. Maxima of total correlation, moving average of the maxima and weighted measure of dispersion for a group of 12 elephants.

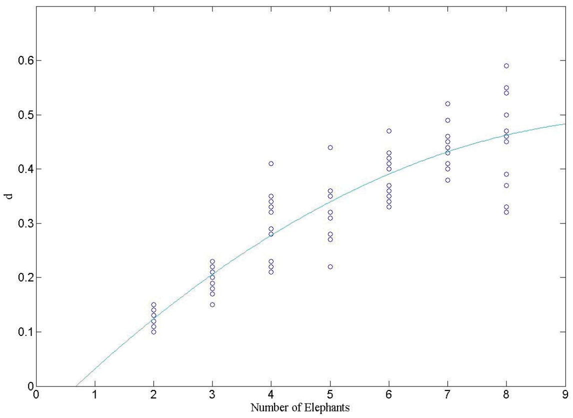

Fig. 10. Dispersion value as a function of size of group for simulated data.

simulation. A monotonic relationship between the values of d and the group size can be seen. The relationship is nonlinear as expected as there is an upper bound on the dispersion. A quadratic fit is used to illustrate this.

VI. CONCLUSIONS

We have proposed a method that functions on several levels. It provides a visual approach to localization and tracking of animals moving along a linear array. Furthermore, our method provides a means to automatically separate our data into segments containing individual passages of either single animals or small herds. Such a segmentation provides information regarding the number of groups of animals passing the array, useful for estimating the sizes of populations of such animals. Finally, the methodology proposed for segmentation, looking at the coherence of the signal at a given time interval relative to noise, also provides information as to the size of the herd. The motivation for the segmentation is to find regions in time for which the signal is relevant across the entire array. This could be used for extraction of individual footsteps to be used for classifying animal type or footstep counting, for example. While this problem is beyond the scope of this paper, the amount of variability in the data as measured by our method seems to offer information as to the size of the herd. As such, it provides important steps towards the goal of censusing elephant populations. Finally, since this method restricts attention to signals that correlate along an angle estimated over a period of time, short bursts of high energy related to "source at infinity" type events will not hinder the method or cause us to think we have observed an additional animal crossing.

REFERENCES

[1] J.D. Wood, C.E. O'Connell-Rodwell and S.L. Klemperer, "Using seismic sensors to detect elephants and other large mammals: a potential census technique," *Journal of Applied Ecology* (42), pp. 587-594, 2005.

[2] J.C. Chen, Y. Kung and R.E. Hudson, "Source localization and beamforming," *IEEE Signal Processing Magazine*, Volume 19, Issue 2, Mar 2002 pp. 30-39, 2002.

[3] A. Savvides, M. Srivastava, L. Girod and D. Estrin, "Localization in sensor networks" Wireless Sensor Networks, pp. 327-349, Kluwer Academic Publishers, Norwell, MA, USA, 2004.

[4] D. Li, K. Wong, Y. Hu and A. Sayeed, "Detection, Classification and Tracking of Targets in Distributed Sensor Networks," *IEEE Signal Processing Magazine*, Volume 19, Issue 2, Mar 2002 pp. 17-29, 2002.

[5] C.H. Knapp and G.C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320-327, 1976.

[6] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk and J. Anderson, "Wireless sensor networks for habitat monitoring," *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, Atlanta, Georgia, USA, Sep. 2002.

[7] S. Gannot and T.G. Dvorkind, "Microphone array speaker localizers using spatial-temporal information," *EURASIP Journal on Applied Signal Processing*, Volume 2006, pp. 1-17, 2006.

[8] H. Teutsch and W. Kellermann, "Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays," Journal of Acoustical Society of America, No. 120 (5), Nov. 2006.

[9] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on a orthonormal decomposition of the soundfield," *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1781-1784, 2002.

[10] I. Potamitis, H. Chen and G. Tremoulis, "Tracking of multiple moving speakers with multiple microphone arrays," *IEEE Transactions on Speech and Audio Processing* 12, pp. 520-529, 2004.