DISCRIMINATION OF BIPEDS FROM QUADRUPEDS USING SEISMIC FOOTSTEP SIGNATURES

Asif Mehmood[1], *Vishal M. Patel*[2] *and Thyagaraju Damarla*[1]

[1]U.S. Army Research Laboratory
2800 Powder Mill Road, Adelphi, MD 20783
[2]Department of Electrical and Computer Engineering
Center for Automation Research, University of Maryland, College Park, MD
email:asif.mehmood1.ctr@mail.mil pvishalm@umiacs.umd.edu thyagaraju.damarla.civ@mail.mil

ABSTRACT

Seismic sensors are widely used to detect moving targets in the ground sensor network, and can be easily employed to discriminate human and quadruped based on their footstep signatures. Because of the complex environmental conditions and the non-stationary nature of the seismic signals, footstep detection and classification is a very challenging problem. The solution to this problem has various applications such as border security, surveillance, perimeter protection and intruder detection. Previous works in the domain of seismic detection of human vs. quadruped have relied on the cadence frequency-based models. However, cadence-based detection alone results in high false alarms. In this paper, we describe a seismic footstep database and present classification results based on support vector machine (SVM). We demonstrate that in addition to applying a good classification algorithm, finding robust features are very important for seismic discrimination.

Index Terms— Seismic signatures, intrusion detection, geophysical signal processing, footstep detection, Wigner-Ville distribution.

1. INTRODUCTION

Personnel detection is an important aspect of intelligence, surveillance, and reconnaissance. It plays a vital role in perimeter and camp protection and in curtailing illegal border crossings [1, 2]. All these applications involve deployment of sensors for a prolonged time; these sensors are often camouflaged so as not to be noticeable by an intruder's visual inspection [3]. Currently, multimodal unattended ground sensors (UGSs) are deployed across our county's border to detect illegal aliens. These UGSs, once deployed, should operate for a prolonged period of time because of their low power consumption. Some of the sensors that require low power are E-field, acoustic, seismic and magnetic. In most UGSs, the imaging sensors are dormant and only wake up to take a picture once the non-imaging sensors determine that there is a viable human target present in the vicinity.

Personnel detection using seismic sensors has been considered by several authors [4, 1, 2] in the literature. Seismic sensors are small enough that they can be easily hidden away so as to not be noticeable from an intruder's visual inspection. Moreover, the creation of artificial vibrations intended to cause confusion in the recognition process is very difficult. Primarily, the seismic sensors are used to estimate the cadence of a person walking [3, 5]. Bland [6] has discussed the use of autoregressive coefficients in designing a footstep detection scheme from acoustic and seismic sensors. Succi et al. [7] proposed the use of signal kurtosis as a test statistic for detection of human footstep signals. Park et al. [8] have considered the problem of detecting and classifying perimeter intrusion using geophones. Iyengar et al. [9] fused acoustic and seismic signals for footstep detection. Their work discusses a novel approach based on canonical correlation analysis and copula theory to establish a likelihood ratio test. Houston and McGaffigan [5] have proposed using cadence features for detection of footsteps.

Most of the cadence-based methods mentioned above are prone to false alarms because humans and animals have similar walking mechanisms and generate similar rhythmic temporal seismic patterns. Furthermore, any quadruped ambling around with a slow cadence can generate the same cadence frequency as the one from a human, or a fast-moving human can generate the same cadence frequency as that of the quadruped. Therefore, it is imperative to come up with robust features and a method that can be used to differentiate quadrupeds from humans based on their walk.

In order to study and develop robust algorithms for discriminating humans from animals using seismic signatures, we have put together a database in which a significant number of footsteps are recorded using geophones under unconstrained outdoor environments. The data is collected using humans and horses. We evaluated a state-of-the-art algorithm on this dataset using support vector machines (SVM) and studied the effect of different features. Based on our limited experiments using this dataset, we make the following observations: extraction of robust features is as important as the recognition algorithms that are used. The performance of recognition algorithms improves gradually as the number of training samples increases. The recognition accuracy varies from low fifties to mid nineties depending on the features and the number of available training data.

This paper is organized as follows: In Section 2, we describe the remote seismic footstep dataset collected by the authors' group. Section 3 describes the classification algorithm along with various features used for classification. We demonstrate experimental results in Section 5 and Section 6 concludes the paper with a brief summary and discussion.

2. DATA COLLECTION

When humans or animals walk, their footsteps generate impulsive seismic signals that propagate through the earth. Seismic signals propagate via body waves (compressional and shear waves) and surface waves (Rayleigh and Love waves). Geophones are used to capture signals generated by these waves. We recorded the seismic data generated by the footsteps of the walking subjects using geophones at a horse farm. The layout of the sensors deployed and the path trajectory of the walking subject is shown in Fig. 1. In Fig. 1, other sensors such as ultrasonic and microphones were also used to collect the data simultaneously. But in this paper, we only study the seismic data. The walking subjects chosen were humans and horses. Primarily, we want to discriminate biped from quadruped on the basis of their footstep signatures. The reason we chose horses in our quadruped category is that they can be easily controlled by the rider. Another reason of choosing horses is that they can generate good seismic data because of their heavy weight. The heavy weight of horses transfer more energy into the ground which results in good seismic signals.

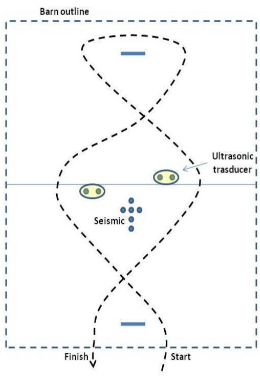


Fig. 1. Walking path and sensors layout in the barn for data collection.

In this paper, we further restricted our work in the discrimination of single walking human form a single walking horse. In our future work, we intend to discriminate multiple people from multiple quadrupeds. In order to capture footstep signatures of a walking human and a walking horse we deployed six seismic sensors as shown in Fig 1. The data was collected with Bruel & Kjaer (B&K) data acquisition (DAQ) system and was saved on the disk for further processing which was done in MATLAB.

3. FEATURE EXTRACTION

Feature extraction is a process of deriving useful information from an original signal, information that is relevant for the task and also has a more compact representation, suitable for use in a classifier. This can be achieved simply through selection, in which elements of the original data vector are kept, or through a transform, which will project the original data in a different, lower-dimensional space. In this paper, we study the following features: short time Fourier transform (STFT), Wigner-Ville distribution (WV), random projection (RP) and linear discriminant analysis (LDA). In what follows, we describe them in details.

3.1. Time-frequency distributions

The footstep signatures of humans or animals obtained using seismic sensors are non-stationary in nature. We employ time-frequency (TF) representations to analyze the non-stationary signals [10]. The goal of TF analysis is to find what frequency occurs at what time in a signal. This means that we have to find a representation for that signal which can relate the time and frequency information. This is performed by mapping a one dimensional signal in the time domain, into a two dimensional TF representation of the signal. The most commonly used method is the STFT, and is defined as follows:

$$STFT(x;t,\omega) = \int x(t+\tau)\; w(\tau)\; e^{-j\omega\tau}\; d\tau \tag{1}$$

where $w(\tau)$ is a sliding window function (e.g., a Hamming window), t is time, and ω is frequency. However, the TF resolution in STFT is not good because of its fixed window size. In order to resolve TF resolution issue, WV distribution is often used. One of the important features of the Wigner-Ville distribution is that it has the best joint TF resolution among all known quadratic joint TF analysis methods [10]. One can compute the WV distribution by applying the fast Fourier transform on the time-dependent autocorrelation as shown in the following equation [11]:

$$WV(x;t,\omega) = \int x(t+\tau/2)\cdot x^*(t-\tau/2)e^{-j\omega\tau}d\tau \tag{2}$$

Despite having high resolution, WV is plagued with the presence of cross-terms interference. The cross-term interference reduces the readability of the TF representation. Because real-valued signals have symmetric positive and negative frequency components, cross-term interference exists between the positive frequency components and the negative frequency components in the WV distribution of real-valued signals. One way to get around this problem is to convert real-valued signals into analytic signals, the cross-term interference between the positive frequency components and the negative frequency components can be suppressed because analytic signals have only the positive frequency components of real-valued signals. In practice, these interference terms can be dramatically reduced by smoothing in time and frequency. A related transform is the smoothed-pseudo Wigner-Ville distribution (SPWVD) which is defined by

$$WV(x;t,\omega) = \int h(t).x(t+\tau/2)\cdot x^*(t-\tau/2)e^{-j\omega\tau}d\tau \tag{3}$$

where $h(t)$ is a smoothing window.

3.2. Linear discriminant analysis

Linear discriminant analysis is a well-known method for feature extraction and dimensionality reduction for pattern recognition and classification tasks [12]. It uses class specific linear methods for dimensionality reduction. It selects projection matrix $\mathbf{A}$ in such a way that the ratio of the between-class scatter and the within-class scatter is maximized. The criterion function is defined as

$$\mathbf{A}_{opt} = \arg\max_{\mathbf{A}} \frac{|\mathbf{A}^T\Sigma_B\mathbf{A}|}{|\mathbf{A}^T\Sigma_W\mathbf{A}|}$$

where $|.|$ denotes determinant of a matrix, Σ_B and Σ_W are between-class and within-class scatter matrices, respectively.

3.3. Random projections

Since we want to embed high dimensional vector into a lower dimensional space, it is important that the relative distances between any two points in the feature space be preserved in the output space. This is characterized by the Johnson-Lindenstrauss (JL) lemma [13], [14]:

Lemma 1 *(Johnson-Lindenstrauss) Let $\epsilon \in (0,1)$ be given. For every set S of $\sharp(S)$ points in $\mathbb{R}^N$, if n is a positive integer such that $n > n_0 = O\left(\frac{\ln(\sharp(S))}{\epsilon^2}\right)$, there exists a Lipschitz mapping $f : R^N \to R^n$ such that*

$$(1-\epsilon)\|\mathbf{u}-\mathbf{v}\|^2 \leq \|f(\mathbf{u})-f(\mathbf{v})\|^2 \leq (1+\epsilon)\|\mathbf{u}-\mathbf{v}\|^2 \quad (4)$$

for all $\mathbf{u}, \mathbf{v} \in S$.

This lemma essentially states that, a set S of points in $\mathbb{R}^N$ can be embedded into a lower-dimensional Euclidean space $\mathbb{R}^n$ such that the pairwise distance of any two points is approximately maintained. In fact, it can be shown that f can be taken as a linear mapping represented by an $n \times N$ matrix Φ whose entries are randomly drawn from certain probability distributions. This in turn implies that it is possible to change the original form of the data and still preserve its statistical characteristics useful for recognition.

4. EXPERIMENTAL RESULTS

The recognition algorithm used in this paper performs LDA followed by a support vector machine [12]. We used a kernel SVM with radial basis function as the kernel function. We evaluate the recognition performance of different features using kernel SVM on the seismic footstep dataset.

4.1. Preprocessing and feature construction

The data was collected simultaneously for all sensors such as seismic, ultrasonics and acoustics, and was sampled at 32 kHz. However, the channels corresponding to seismic signals on DAQ were later downsampled to 1 kHz for faster computation. In preprocessing stage, the portion of data that contained footstep signatures with reasonably significant energy are considered. As the data was collected for the whole run which lasted for 60-90 seconds, only 15-20 seconds of the portion contained the footsteps. The time domain (TD) data is shown in Fig. 2. It can be seen in Fig. 2 (a) and (b) that when a subject approaches the sensors the signal level increases and as the subject moves away from the sensors the signal level decreases. An enlarged portion of Fig. 2 (a) and (b) is shown in Fig. 2 (c) and (d), respectively.

Once the data has been preprocessed, TF distributions are used to extract the joint time-frequency information present in the signal. Fig. 3, shows examples of Wigner-Ville distributions corresponding to human and horse footstep signatures. Note that one can clearly see the difference between the TF representation of human and horse footsteps.

After the TF distributions are obtained, either LDA of RPs are performed to transfer the data into low dimensional space. Fig. 4 (a) shows the scatter plot corresponding to the TF distribution of the human and horse data. Fig. 4 (b) shows the scatter plot after the linear discrimination operation. From Fig. 4 (b), we see that the spectral components of human footstep signatures and the spectral components of horse footstep signatures are adequately separated. This separated data is then supplied to an SVM for classification.

4.2. Classification results

We consider the following six different types of features:

- LDA on Time domain signals (LDA-TD)
- RPs on Time domain signals (RPA-TD)

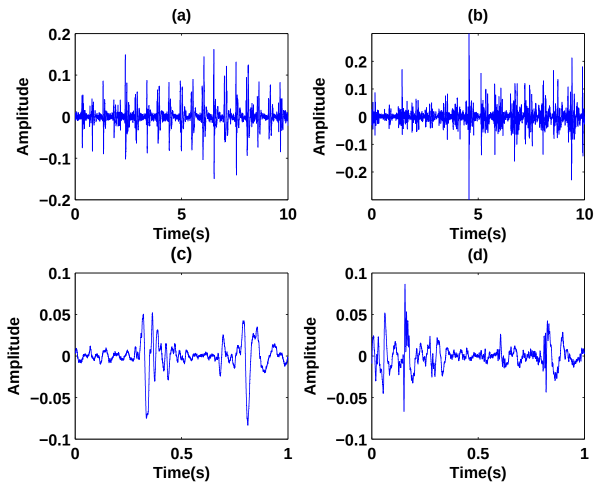


Fig. 2. Raw data: (a) Human footstep signatures, (b) Horse footstep signatures, (c) Enlarged human footstep signatures, (d) Enlarged horse footstep signatures.

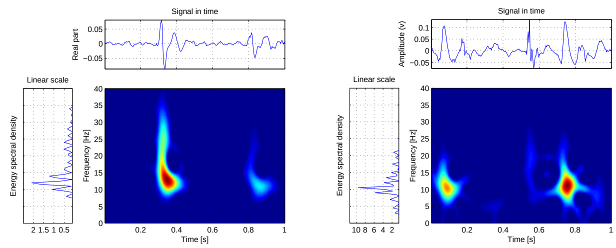
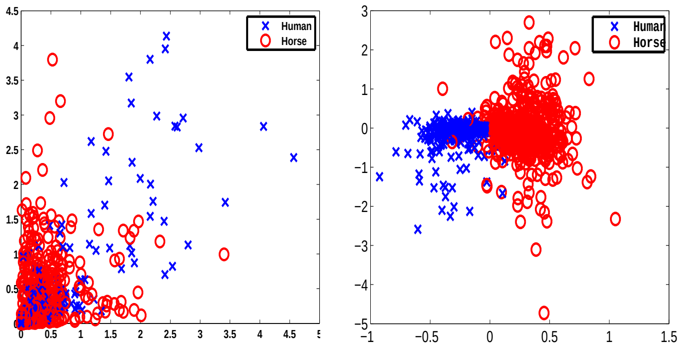


Fig. 3. Wigner-Ville distributions. Left: Human footstep signatures. Right: Horse footstep signatures.



(a) Before LDA. (b) After LDA.

Fig. 4. Linear discriminant analysis of human and animal signatures.

Table 1. Classification rates (CR) (in %) corresponding to different features.

Feature	LDA-TD	RPA-TD	LDA-FD	RPA-FD	LDA-WV	RPA-WV
CR	76.67	74.13	86.93	84.24	**92.05**	85.45

- LDA on short time Fourier transformed signals (LDA-FD)
- RPs on short time Fourier transformed signals (RPA-FD)
- LDA on Wigner-Ville transformed signals (LDA-WV)
- RPs on Wigner-Ville transformed signals (RPA-WV).

Classification is performed with different number of training samples. The number of test samples were kept fixed at 250. Fig. 5 shows the performance of different features using kernel SVM on this dataset. It is clearly seen that with increase in number of training sample, the classification rate (CR) for each scenario increases. It is evident from the figure that time domain features both randomly projected and linearly discriminated, performed poorly. However, when the data was transformed into its spectral components and features were obtained using RPs and LDA, then the correct classification rate was improved significantly. Both the random projections and LDA performed well, but the best results were obtained when LDA was performed on the Wigner-Ville transformed signals. With 100 training vectors and 250 test vectors, the CR=92.0% is achieved for WV-LDA case. The best classification rates for each feature type were achieved using 100 feature vectors and are reported in Table 1.

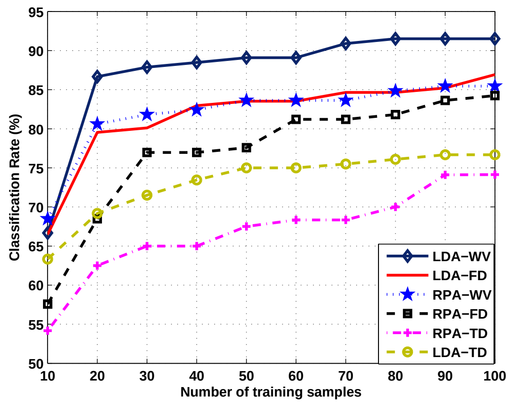


Fig. 5. Classification rate vs. number of training samples.

It is worth mentioning here that the test data and training data were collected at different times at different sites. We also tested our algorithm on the data that contained both training and test data from the same site and collected at the same time. In this case the best recognition rate was found to be 98.0%. This is because there was no significant variability in the data.

5. DISCUSSION AND CONCLUSION

In this paper we presented a method of footstep signature analysis for human and quadruped (horse) discrimination using seismic sensors. Different features of the footstep signatures are used for SVM classifier. It has been shown that human footstep signatures can be distinguished from quadruped footstep signatures utilizing the proposed method. It has also been shown that the proposed method lowers the false alarm rates and improve classification capabilities. The results proved that the proposed method can be used for practical situations such as protecting military assets and troop bases from approaching unauthorized personnel and protecting borders from illegal aliens and drug traffickers.

In recent years there has been a great interest in applying sparse representation-based and dictionary learning-based approached for classification. In our future work, we plan to apply these techniques for classification problems discussed in this paper. We also plan to collect data from quadruped other than horses and test our approaches in human classification vs. animals. We also intend to classify genders based on their footstep signatures.

6. REFERENCES

[1] T. Damarla and D. Ufford, “Personnel detection using ground sensors,” in *Proc. of SPIE*, Orlando, FL, 2007, vol. 656205, pp. 1–10.

[2] H. Park, A. Dibazar, and T. Berger, “Protecting military perimeters from approaching human and vehicle using biologically realistic dynamic synapse neural network,” in *Technologies for Homeland Security, 2008 IEEE Conference on*, Boston, MA, May 2008, pp. 73–78.

[3] H. Park, A. Dibazar, and T. Berger, “Cadence analysis of temporal gait patterns for seismic discrimination between human and quadruped footsteps,” in *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, Taipei, Taiwan, 2009, pp. 1749–1752.

[4] S. Schumer, “Analysis of human footsteps utilizing multi-axial seismic fusion,” in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 697 –700.

[5] K. M. Houston and D. P. McGaffigan, “Spectrum analysis techniques for personnel detection using seismic sensors,” in *Proc. SPIE*, Orlando, FL, 2003, vol. 5090, pp. 162–173.

[6] R. Bland, “Acoustic and seismic signal processing for footsetp detection,” M.S. thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, Cambridge, MA, 2006.

[7] G. Succi, D. Clapp, R. Gampert, and G. Prado, “Footstep detection and tracking,” in *Proc. of SPIE*, Orlando, FL, 2001, vol. 4393, pp. 22–29.

[8] H. Park, A. Dibazar, and T. Berger, “The application of dynamic synapse neural networks on footstep and vehicle recognition,” in *Proc. International Joint Conference on Neural Networks IJCNN*, Atlanta, GA, 2007, pp. 1842–1846.

[9] S. G. Iyengar, P. K. Varshney, and T. Damarla, “On the detection of footsteps based on acoustic and seismic sensing,” in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, Pacific Grove, CA, November 2007, pp. 2248 –2252.

[10] P. Goncalves F. Auger, P. Flandrin and O. Lemoine, *Reference Guide to: TimeFrequency Toolbox for Use With MATLAB*, 1995.

[11] P. Martin, W.; Flandrin, “Wigner-ville spectral analysis of nonstationary processes,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 33, no. 6, pp. 1461 – 1470, 1983.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, USA, second edition, 2001.

[13] W. B. Johnson and J. Lindenstrauss, “Extensions of lipschitz maps into a hilbert space,” *Contemp. Math.*, pp. 189–206, 1984.

[14] D. Achlioptas, “Database-friendly random projections,” *ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems*, pp. 274–281, 2001.