CADENCE ANALYSIS OF TEMPORAL GAIT PATTERNS FOR SEISMIC DISCRIMINATION BETWEEN HUMAN AND QUADRUPED FOOTSTEPS

Hyung O. Park, Alireza A. Dibazar, and Theodore W. Berger

Biomedical Engineering Department of the University of Southern California
1042 Downey way, DRB140, Los Angeles, CA 90089

ABSTRACT

This paper reports on a method of cadence analysis for the discrimination between human and quadruped using a cheap seismic sensor. Previous works in the domain of seismic detection of human vs. quadruped have relied on the fundamental gait frequency. Slow movement of quadrupeds can generate the same fundamental gait frequency as human footsteps therefore causing the recognizer to be confused when quadruped are ambling around the sensor. Here we propose utilizing the cadence analysis of temporal gait pattern which provides information on temporal distribution of the gait beats. We also propose a robust method of extracting temporal gait patterns. Features extracted from gait patterns are modeled with optimum number of Gaussian Mixture Models (GMMs). The performance of the system during the test for discriminating between horse, dog, multiple people walk, and single human walk/run was over 95%.

Index Terms — Cadence analysis, pattern classification, feature extraction

I. INTRODUCTION

With the growing interest on security problems, the development of technologies that can detect potential threats such as a human or vehicle approaching military assets has been stimulated. One area of interest is to utilize seismic waves propagating from the source i.e. a threat in order to recognize the threat. 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.

The signal measured from a geophone has a 0.1Hz~100Hz frequency range due to the resonant characteristics of the sensors. Although the frequency response of the seismic sensor is in a narrow frequency band, spectral analysis can be used for discriminating between seismic events caused by human footsteps (or four-leg animals) and vehicles. However, due to the very similar walking mechanism of humans and animals, the generated rhythmic temporal seismic patterns of humans and animals are very similar. This renders the discrimination between a human's and an animal's footstep using frequency analysis a failure.

We have already shown that a biologically realistic neural network which captures short-term dynamics of the signals can be employed for the seismic recognition of human footsteps and vehicles [1]. With analysis of the seismic recordings of four-leg animals' footsteps, we also realized the importance of long-term features for discriminating between four-leg animals' and human footsteps. Obviously, cadence analysis is a good candidate for detecting human presence in any situation since there is no seismic source generating the frequency found in human gaits even in urban areas.

The work by Succi et al or Houston and McGaffigan has proposed utilizing cadence features for seismic detection of footsteps using the geophone sensor [2-3]. Moreover, a few other works also have employed cadence analysis using different sensors [4-5]; all of the mentioned research have used the fundamental gait frequency as a main feature. However, any quadruped ambling around with a slow gait can generate the same gait frequency as the one from human. Therefore, we propose utilizing temporal patterns of gait period as an additional feature to the gait frequency. In addition, employing the temporal pattern of the gait enables us to discriminate between multiple people footsteps and a walking horse. The latest can not be done if only cadence analysis be performed.

In this paper, we only focus on temporal pattern analysis of seismic events i.e. cadence analysis to see how statistically different the patterns of human and quadruped are and to check the possibility of cadence analysis as one of features for footstep recognizer.

II. METHODS

A. Feature extraction

The goal of this section is to introduce the features and feature extraction method for cadence analysis of temporal gait patterns (Figure 1).

After applying a three second sliding window – with two seconds overlap – on the incoming signal, the signal is passed through a band-pass filter to enhance the Signal to Noise Ratio (SNR). Then a Hilbert transform and low pass filtering (smoothing process) is applied to extract the envelope of the signal.

In the next step, we utilize this signal to extract the mean temporal pattern of the gait by averaging over each gait periods. Therefore it is necessary to estimate the gait period within the

This project was supported in part by grants from ONR, NAVY, and DARPA. Authors of this manuscript would also like to thank Safety Dynamics Inc. for their support and all members of USC's Neural Dynamics Laboratory especially Sageev George for their invaluable help.

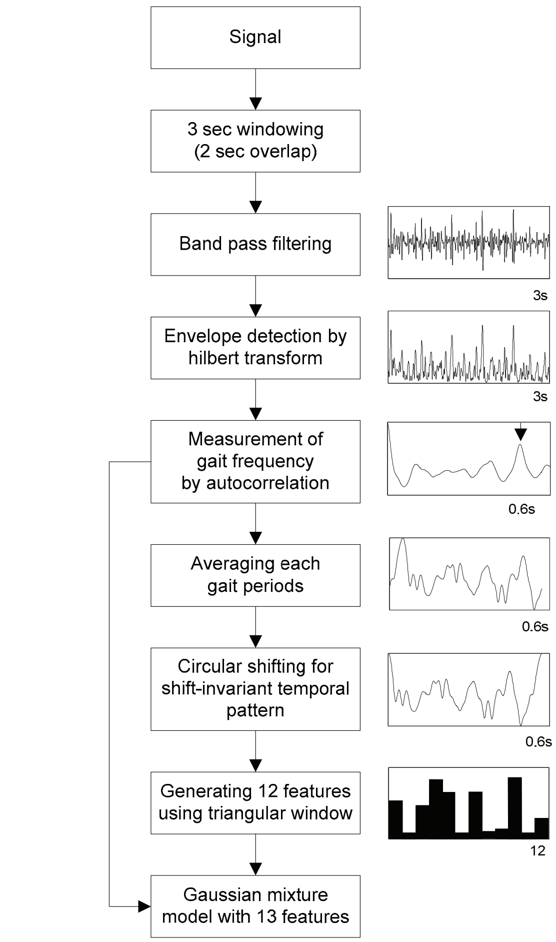

Fig. 1: Feature extraction algorithm and examples of the output of each block

three second window and partition the three seconds signal based on gait period. This can be achieved in the following steps:

A) Gait period is estimated by using the auto-correlation function. Because of the periodicity in the signal auto-correlation signal has local maxima at the time of gait period. In general, finding the local maxima is a challenge, however, due to the resonant characteristics of the seismic sensors and the periodicity from walking mechanism, there is a detectable peak in the auto-correlation function. It also worth to mention that the gait period (or cadence frequency) will later be employed as one the features.

B) Using the estimated gait period in A, the three second window is equally divided into k number smaller windows each having gait period length.

C) The partitioned signals from B are averaged.

D) In order to make a shift-invariant temporal gait pattern representation, the averaged gait pattern from C is circular-shifted so that the local maxima of the pattern is on the first sample. The partitioning of the three second signal into *k* frames will have some remainder which is considered in the circular shift of the next consecutive frame.

Lastly, 12 triangular weighting functions are applied to the temporal pattern acquired from the steps explained above so that the gait temporal pattern can be represented by 12 features. Considering the feature obtained from step A, the total number of features is 13.

B. GMMs

The Gaussian Mixture Models (GMMs) is one of the most well-known and useful classifiers, having been widely used in many applications. For a multimodal random variable, whose values are generated by one of several independent sources, a finite mixture model can be used to approximate the true probability density function. Moreover, GMM is a good candidate as a classifier when there exists no prior knowledge of a probability density function. Therefore, estimating the distribution with GMM not only provides a chance to have a general model but also helps to understand the phenomena for a better use of the information of the distribution.

A non-singular multivariate normal distribution of a D dimensional random variable $X \leftrightarrow x$ can be defined as

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \quad (1)$$

where μ is the mean vector and Σ the covariance matrix of the normally distributed random variable X.

The GMM can be defined as a weighted sum of Gaussians function:

$$p(\boldsymbol{x};\boldsymbol{\theta}) = \sum_{c=1}^{C} \alpha_c \mathcal{N}(\boldsymbol{x};\boldsymbol{\mu}_c,\Sigma_c) \quad (2)$$

where α_c is the weight of c^{th} mixture and θ is defined as following.

$$\boldsymbol{\theta} = \{\alpha_1,\boldsymbol{\mu}_1,\Sigma_1,\ldots,\alpha_C,\boldsymbol{\mu}_C,\Sigma_C\} \quad (3)$$

To estimate or train the model parameter θ, the Figueiredo-Jain (FJ) algorithm was used [6], which automatically chooses the optimum number of mixtures during the training. The objective function of this algorithm utilizes the minimum message length criterion for finding optimum number of mixtures as defined in the equation (4) so that it can select best model directly from data rather than hierarchy of model-class.

$$\Lambda(\boldsymbol{\theta},\mathsf{X}) = \frac{V}{2}\sum_{c:\,\alpha_c>0} \ln(\frac{N\alpha_c}{12}) + \frac{C_{\mathrm{nz}}}{2}\ln\frac{N}{12} + \frac{C_{\mathrm{nz}}(V+1)}{2} - \ln\mathcal{L}(\mathsf{X},\boldsymbol{\theta}) \quad (4)$$

where N is the number of training points, V is the number of free parameters specifying a component, and C_{nz} is the number of components with nonzero weight in the mixture ($\alpha_c > 0$). The last term $\ln\mathcal{L}(\mathsf{X},\boldsymbol{\theta})$ is the log-likelihood of the training data given the distribution parameters θ.

III. Experiment and Results

A. Data recording

A horse was chosen for quadruped class because the gait can be easily controlled by a rider and also data can be easily acquired with a rider's control. In addition, the signal itself is clearer than that of a dog due to the high energy transferred from its weight. From a horse ranch of Yucca Valley, CA, a 9 year-old Hawaiian mustang was recorded using a geophone, a low-cost seismic sensor and developed hardware unit, at an arena and a hill

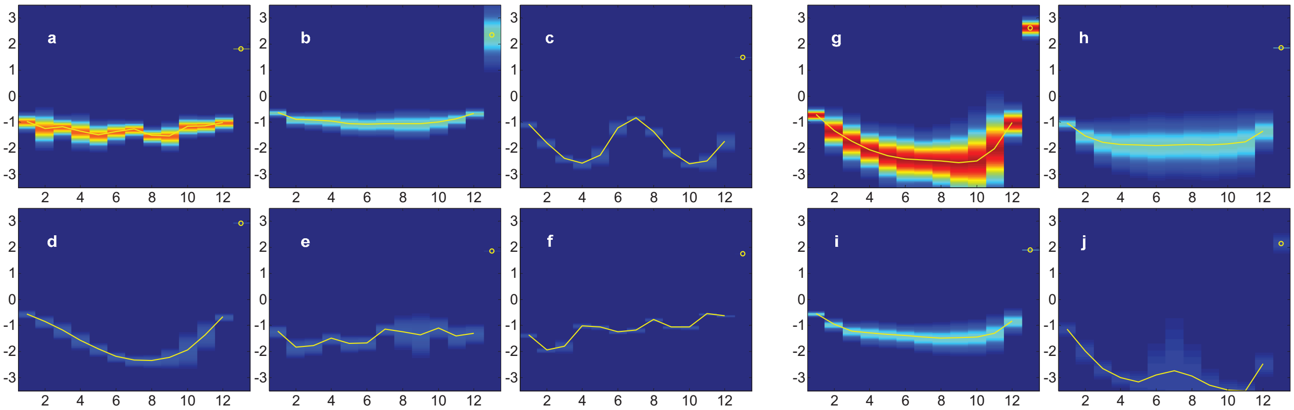

Fig. 2: Feature vectors for horse and human classes; each plot represents an independent Gaussian mixture. X-axis is the feature number (1^{st} ~ 12^{th}: cadence pattern, 13^{th}: gait frequency) and y-axis is normalized amplitude for the 1^{st} ~ 12^{th} features and frequency for the 13^{th} feature. 6 Gaussian mixtures from a to f for the horse, 4 Gaussian mixtures from g to j for the human were built from the training data set. The yellow lines for the 1st~12th features and the yellow circles for the 13th feature are the mean value of each feature and the color code represents its distribution.

in the early morning. First, recordings were made as it walked and ran around the arena with different gaits e.g. in order of speed, walk; 4-beat gait, trot; 2-beat gait, and canter; 3-beat gait for 20 minutes keeping a distance of maximum 100 feet from the sensor. The recorded data also includes a different type of walk which is called collective walk or working walk and the transition gait between each gait, which is none of the above natural gaits. The canter gaits appeared only in short periods mixed with the walking gait and mostly slow canter which was slower than trot. Second, at the hill, the data of gallop, which is the fastest 4-beat gait, and the other natural gaits were recorded for another 20 minutes of walking and running around the hill. The distance from the sensor was from 20 feet to 200 feet.

For human footsteps, the data of a single person running and people – group of five – walking in a group were collected at a sandy terrain near the Joshua Tree national park, CA again using the geophone sensor. Each of 4 different people ran along a straight path of 200 feet and data was recorded for over 5 roundtrips with speed varying from the fast running speed possible down to fast walking. For the data of people walking in a group, 5 people walked naturally along the same path in a group for 5 roundtrips. Then, the same 5 individuals were recorded walking at the same rate of speed and in sequence, keeping 6 feet from person to person, for another 5 roundtrips. Also, they were recorded walking randomly around the sensor for 3 minutes. The sensor was located 5 feet away from the middle of the path.

B. Training

After preprocessing of the data, only human and quadruped's footsteps were detected from the input signal and the other classes were rejected. The rejected data includes background i.e. no event, any event with no gait frequency in our interested frequency band, and transition in speed and gait pattern. The preprocessing includes filtering at 10~100Hz and applying a threshold to the auto-correlation function at a window corresponding to 1.4Hz~7Hz gait frequency. Features discussed in section II.A extracted from pre-processed data and GMMs were setup to model the features.

As a result of equation (4), 6 Gaussian mixtures for the horse, and 4 Gaussian mixtures for the human classes were formed during training process. The mean value and the distribution of each mixture are presented in the Figure 2.

Figures *2.a* to *2.f* presents the statistics of horse's cadence pattern trained by mixtures. The mixture shown in Fig. 2.*a* is the most likely pattern in the data set for detecting horse and the others (*2.b* to *2.f*) are presented in the order of their generating likelihood. The mixture shown in Fig. 2.*a* represents also "walk" which is a 4-beat gait. The mixture *e* and *f* are representatives of the other types of the "walk" gait (all of the "walk" gaits show 4 peaks on their temporal patterns). The mixture *e* includes the pattern of slow canter which is slow 3-beat gait and in general the feature number 1, 7, and 10 represent the peaks of 3 beats. The mixture *b* represents the gallop which is the fastest 4 beat. In Gallop, the peaks were not observed due to relatively higher variation of the location of the peaks in time and shorter duration of their time period. The mixture *c* and *d* are built for trot which is a fast 2-beat gait. Similarity between two time domain peaks has doubled the gait frequency in the mixture *d*.

Figures *2.g* to *2.j* show the mixtures of human cadence pattern. The mixture *g* is the most likely pattern for human which is built from a single person's footsteps including running and walking. Although human gait is 2-beat, most of human footsteps have the similarity between two 2-beats footsteps so that the gait frequency is measured doubled as in the mixture *g*. The mixture *h* and *i* represents multiple people's walk. Randomness of the location of peaks in time made the feature space to be flat and the personal variance of the strength of footstep provides the difference between the 1st feature and the others.

C. Results

To evaluate the performance of the trained recognizers, we performed self-validation test on the data we used for training. During the test, the average of posteriori probabilities of each class on 10 consecutive window frames was calculated (we have assumed that there is no abrupt changes within the class). We found that the average posteriori is a powerful technique in our application which enhances the low-SNR observations results and reduces false positives.

Sample test signals and their results are plotted in the Figures 3 and 4. Also the classification results of the experimants are summarized in Table I. For the data set of each class, the number of frames with wrong recognition was counted and its percentage is presented. The test was also conducted separately too see if the method of the study can discriminate between multiple people (5 persons) walking, running, and horse. We also tested the system with human (a single person) walking data which we did not train the system with and had recorded a year ago at the same place.

Based on Table I, higher false recognition rate on human running arises from the similarity to the trot gait of the horse. On the specific gait frequency, a human's cadence pattern and a horse's are very similar as can be seen in Figure 3 and 4, also in the Figure 2.*d* and 2.*g*. However, overall performance as shown in the Table I is over 95% correct recognition.

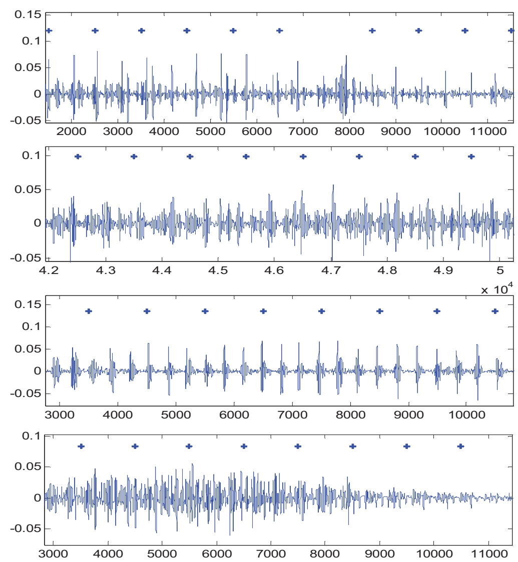

Fig. 3: Example of horse's typical footsteps and their recognition: From top to bottom, each plot represents the temporal signal from walk, canter, trot, and gallop respectively. X-axis is sample in 1/1000s. Note blue crosses on top of the signal meaning recognized as horse's footstep.

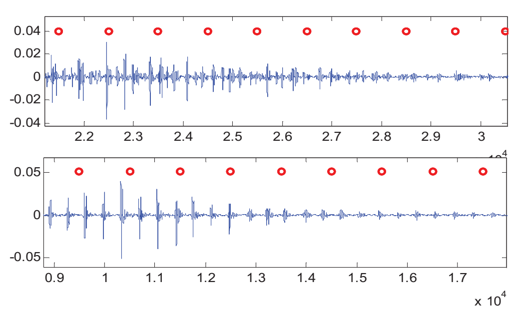

Fig. 4: Example of human's typical footsteps and their recognition: The top plot represents the temporal signal from multiple people walking, and the bottom one running. Note red circles on top of the signal meaning recognized as human's footstep.

TABLE I. FALSE RECOGNITION RATE FOR HUMAN AND HORSE

Test set	False recog. (%)		Total frames	
	Human	Horse	Human	Horse
People walk	1.98		553	
Human run	5.19		617	
Horse		1.86		1561
Human walk (single person)	1.46		3222	

IV. DISCUSSION AND CONCLUSION

In this paper we presented a method of cadence analysis for human and quadruped discrimination using vibration sensors. The fundamental gait frequency and temporal pattern of gait was used as features for GMM classifier. A horse was chosen for quadruped and it was shown that human gait can be distinguished from quadruped gait utilizing the proposed method. Although we did not present the result from dog, dog's gait was also recognized as quadruped without any additional training suggesting that the model trained with horse can be an appropriate representative for quadruped. In addition, after investigation of the test results, the cadence analysis turned out to be mutually complementary to short-term analysis of the seismic events presented in [1]. Therefore, the integration of short-term and cadence analysis to detect other classes of security breaches (e.g. seismic signals generated by small unmanned, and heavy track vehicles) will be an immediate next step.

REFERENCES

[1] Hyung O Park; Dibazar, A.A.; Berger, T.W., "Protecting Military Perimeters from Approaching Human and Vehicle Using Biologically Realistic Dynamic Synapse Neural Network," *Technologies for Homeland Security, 2008 IEEE Conference on* , vol., no., pp.73-78, 12-13 May 2008

[2] G. Succi, D. Clapp, and R. Gambert, "Footstep, Detection and Tracking," in *Proc. of the SPIE, Unattended Ground Sensor Technologies and Applications III*, 2001, SPIE vol. 4393, pp. 22-29.

[3] K. M. Houston and D. P. McGaffigan, "Spectrum Analysis Techniques for Personnel Detection using Seismic Sensors," in *Proc. SPIE Symposium on Unattended Ground Sensor Technologies and Applications V*, 2003, SPIE vol. 5090, pp. 162-173.

[4] Sabatier, J.M.; Ekimov, A.E., "A Review of Human Signatures in Urban Environments Using Seismic and Acoustic Methods," *Technologies for Homeland Security, 2008 IEEE Conference on* , vol., no., pp.215-220, 12-13 May 2008

[5] Zhaonian Zhang; Pouliquen, P.; Waxman, A.; Andreou, A.G., "Acoustic Micro-Doppler Gait Signatures of Humans and Animals," *Information Sciences and Systems, 2007. CISS '07. 41st Annual Conference on* , vol., no., pp.627-630, 14-16 March 2007

[6] M. Figueiredo and A. Jain, "Unsupervised learning of finite mixture models," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 3, pp. 381–396, 2002.