Extracting clasification features from seismic sources

Varbanov Vladimir
Institute of Metal Science, Equipment and Technology with Hydroaerodynamic Center at Bulgarian Academy of Sciences, 67 Shipchenski Prohod Street, 1574 Sofia, Bulgaria
vdvarbanov@abv.bg

Abstract: *Automatic classification of seismic sources finds vast application security domain. Due to restricted computational power in the edge devices a robust and less complex features need to be extracted and sent to appropriate classifier. Such features are histograms of oriented gradients. It represented the relative spectral distribution of derivative of amplitude against frequency instead of average spectral envelope.*
KEYWORDS*: SVM (SUPPORT VECTOR MACHINES), HOG (HISTOGRAM OF ORIENTED GRADIENTS) DESCRIPTOR, MACHINE LEARNING*

1. Introduction

With advancement of computer vision many researchers try to implement robust algorithms and feature descriptors of this domain [1] into classifying seismic signals so the make spectrograms and of the signal with respect of time and try to classify the objects as if they are image. In example [1] and [2] use convolution neural networks, while [4] use local binary descriptors, and [5] use wavelets. Although the results are impressive there are several factor making this classification features and classifiers hard to implement in real time classification working on the edge due to the following reasons:

- lack of large open access database with seismic signals. The fairly small training database makes the use of neural networks unpractical;
- training neural networks require more time, because of the high computational complexity of the training process, and hardware with at least average GPU is not available to the researcher;
- extracting features based on wavelet analysis are computationally expensive and can not be applied on low energy consumption microcontroller, supposed to work on the field for prolonged amount of time, because such microcontrollers lack the computational power to perform extensive calculations like multiple convolutions of signal with different wavelet functions;
- applying computer vision algorithms directly needs several measurements of the signal and concatenating them as image, which requires not only memory, but also significantly more time in order to accumulate these measurements. For moving vehicles by the time enough data is collected the target may be long gone. Even if we sample the signal in shorter time frames, the time/frequency resolution ambiguity which is typical for the Fourier transformation will take effect of the resolution of the specter, and because of the very narrow frequency range of such signals – occurring between 1[Hz] and 300[Hz] time frame length for acquiring the measurement is of high important, improving the frequency resolution by adding “buckets of zeros” works only to a certain extent.

Taking into consideration the specifics of the task – training classifier with small database, in narrow frequency bandwidth on hardware with low computational resource, supposed to perform on low power consumption/low processing power device, the most practical approach is to use traditional machine learning algorithms like Gauss Mixture Model(GMM), Support Vector Machine(SVM), Principle Component Analysis(PCA), Linear Discriminant Analysis and/or other types of classifiers. The classifier of choice Linear Support Vector Machine in which the data can be classified into two linearly separated classes. In the base form, linear separation, SVM tries to find a line that maximizes the separation between a two-class data set of 2-dimensional space points. To generalize, the objective is to find a hyperplane that maximizes the separation of the data points to their potential classes in an n-dimensional space. The data points with the minimum distance to the hyperplane (closest points) are called *Support Vectors*. The hyper line takes the form of line if classification is based on 2 features, a plane for 3 features or volume figure for 4 and more dimensions of the features used to characterize the objects being classified. The current paper use binary SVM classification and is focused on extracting robust features to be fed to the classifier.

2. Feature extraction

Seismic signal sources used in current paper are wheeled and tracked vehicles, more specifically:

- bat-2m, which is bulldozer with tank chassis, refered further in paper as “bat”.
- kraz_tmm, six wheeled truck designed to carry, pontoons, referred in the paper as “kraz_tmm”;

Features used for classification are taken from the frequency domain of the signals in order to achieve phase invariance. Unlike neural network, where features are extracted automatically based on the large training dataset, here the features must be handcrafted due to small volume of data. Therefore a spectral representation of the signal need to be made, as shown in figures below[7,8]. The data is taken on 22.06.2020 and 23.06.2020 with sensor with id: 1.1, sensitive in frequency range between 1[Hz] and 300[Hz].Two runs are made for every vehicle:

- forward -vehicle moving away from the sensor;
- backward – vehicle moving towards the sensor.

Fig. 1 and Fig. 2:

After fft transformation with hamming window and windows overlap 0.75 the following spectrograms are constructed:

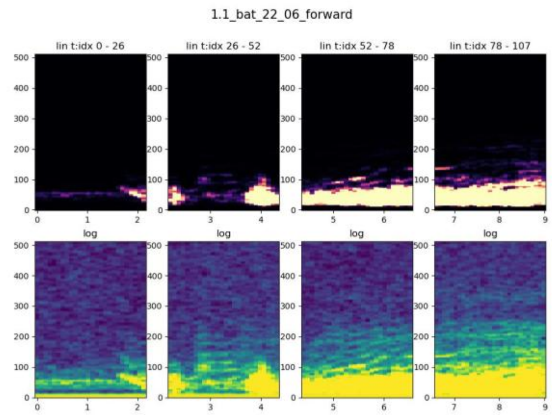

Fig. 1 *Spectrogram for bat_2m “bat”, moving away from the sensor.*

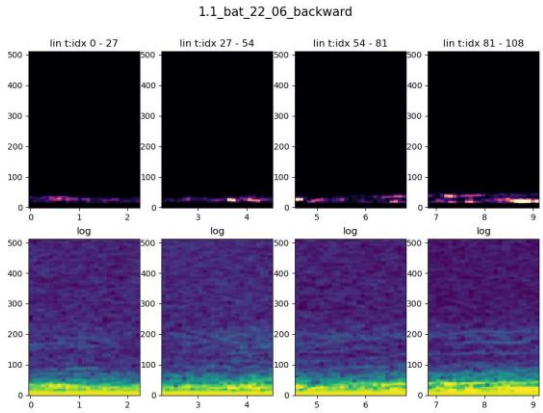

Fig. 2 *Spectrogram for bat_2m "bat", moving towards the sensor.*

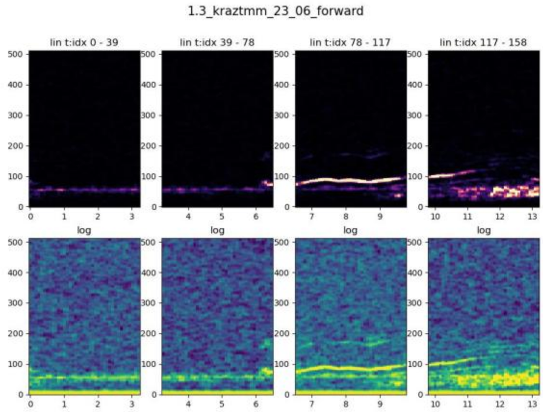

Fig. 3 Spectrogram for bat_2m "bat", moving away from the sensor.

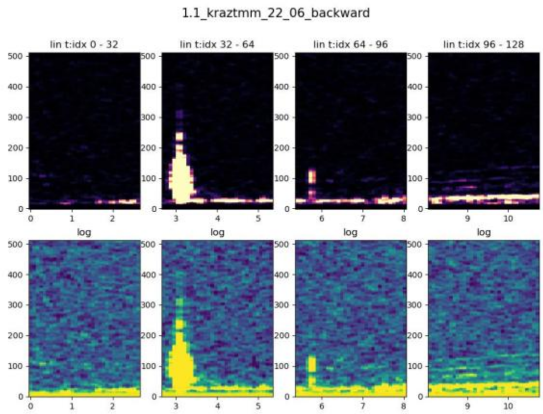

Fig. 4 *Spectrogram for bat_2m "bat", moving away from the sensor.*

The figures show spectrograms in linear and logarithmic scale. Where the linear scale is on the first row, and logarithmic scale on the second. Vertical axis shows the frequency in Hertz and horizontal axis shows the time in seconds. Each spectrogram is divided in to four columns for better scaling of the amplitudes of the signal which are coded as color. Although signal is not processed as image for the reasons stated in introduction, it is still useful to observe the whole seismic picture of the process. For example frequency shift is observed in last two columns in fig. 1, fig.3, fig.4. The effect happens at the end of the run when the vehicle has gained speed. This effect is particularly visible in fig.4 row 2, columns 3 and for where the peak at the 5.4-th second is observed at frequency 89[Hz] and reaches frequency 145[Hz] in the end of the recording at the 10.8-th second. Therefore simple frequency matching of the peaks would not be sufficient as feature descriptor. Although it could be used as feature with lower weight on the decision in future work. A robust feature must take into account the curvature of the spectral distribution for single frame, and allow for frequency and amplitude fluctuations of the signal[4]. If such feature is based on histograms in the sub bands of the spectrum it could compensate for said fluctuations. Another consideration is the selection of right amount granularity. What is meant by granularity is the width bins of histogram, used as classifying feature. Such feature is the histogram of oriented gradients, used in computer vision. In general this features works by finding contours in an image and creating histogram of this gradient's magnitude g and angle θ. In our case the HOG descriptor has to be adapted to one dimensional case. The steps for feature extracting are: HOG is implemented in the following steps:

- creating series of Fourier samples on the signal temporal frequency spectrogram as shown in fig. 5 and fig. 6:

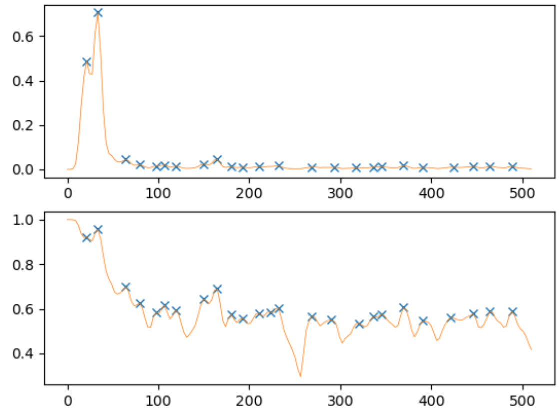

Fig. 5 *Specter of bat_2m "bat", moving towards the sensor at the 0-th second in linear (upper) and logarithmic (lower) scale. Frequencies are displayed on horizontal axis, normalized amplitudes are displayed on the vertical axis.*

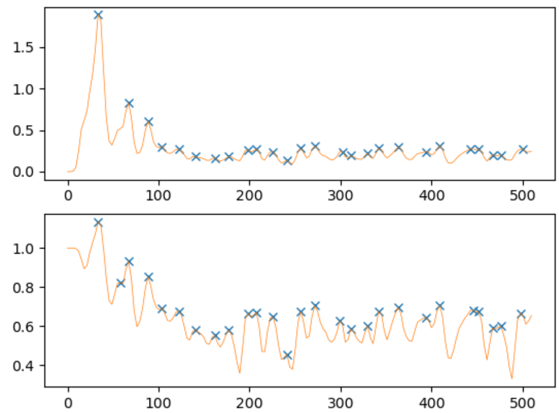

Fig. 6 *Specter of kraz_tmm "kraz_kraz", moving towards the sensor at the 8.60064-th second in linear (upper) and logarithmic (lower) scale. Frequencies are displayed on horizontal axis, normalized amplitudes are displayed on the vertical axis.*

For further processing only the logarithmic scale is to be used as it has better amplitude range

- processing the spectrogram like an image by
- applying gradient filter for the vertical component of the signal $-1\ \ 0\ \ 1$. This filter moves

horizontally through the signal and in current "position" of the signal writes the difference between the next and the previous value of the amplitude.

- Finding the absolute magnitude of the gradient the angles gradient in every position of the spectrogram using formulas (1) and (2)

$$g_i = abs(S_{i+1} - S_{i-1}), \quad (1)$$

where S_i is the amplitude at the i-th frequency

$$\theta_i = arctg(g_i/\Delta f) \quad (2)$$

where Δf is the difference between the i-1–th and i-+1-th frequency.

- Split the spectrogram into cells with suitable length, which is chosen heuristically (in this paper the chosen cell size is 12). For every cell perform calculations based on formulas 1 and 2. Classical HOG splits spectrogram into blocks, which then are split further into cells with overlapping windows, such procedures are not done here, as they do not improve the accuracy of the classifier.
- Create a vector of length 9 with where every position represents the slope of the signal in current cell, in degrees between 0 and 180 with step 20 degrees. The value in each position in the vector represent the sum of the amplitudes in cell occurring in the slope bin.
- Normalize each vector and concatenate

The resulting hog descriptors extracted from the logarithmic scales specters in fig.5 and fig.6 are shown in fig. 7.

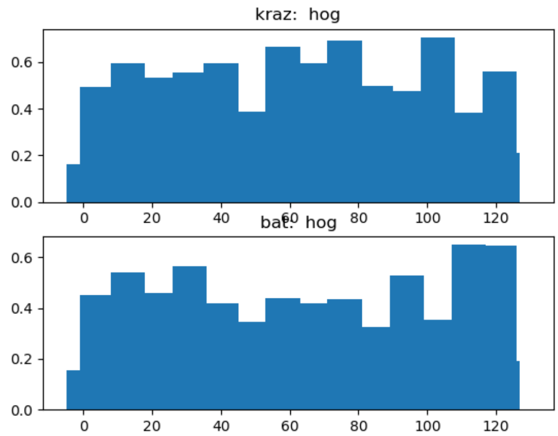

Fig. 7 *HOG descriptor for kraz_tmm and bat. The vertical axis represents the normalized values calculated in step 4 from the procedure, the horizontal value represent the number of features extracted from single fft frame.*

3. Classification results

In order to evaluate the robustness of the features we extracted from the signal, we train linear SVM classifier with 75% of the samples we got from the spectrograms figures 1 to 4. The other 25% will be used for testing. The general idea is that the classifier should be tested on samples it "sees" for first time. In order to visualize the results we use confusion matrix. On the vertical scale are shown the true signal source of each sample belongs, and on the horizontal the predicted ones.

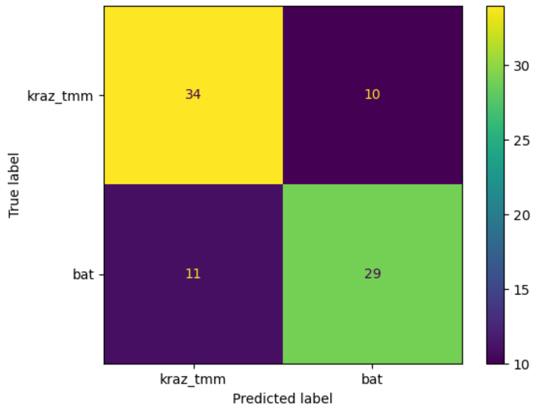

Fig. 8 *confusion matrix of binary SVM classifier with clases kraz_tmm and bat.*

The volume of testing data set of kraz_tmm consisted of 44 samples, while the data set of bat consisted of 40 samples. From fig. 7 can be seen that 34 of 44 kraz_tmm samples are correctly recognized, while 10 has been misclassified as bat. Regarding the other class 29 of 40 samples has been correctly recognized as bat and 11 has been misclassified. In order to get the chance of correct prediction of each class we normalized values in fig. 8 for each row. The result is shown in fig. 9.

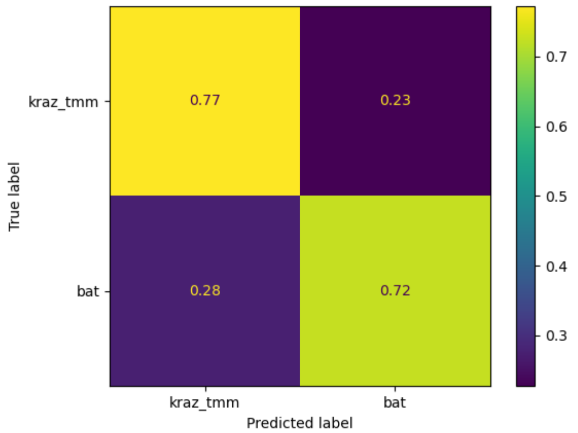

Fig. 9 *normalized confusion matrix of binary SVM classifier with clases kraz_tmm and bat.*

As it can be seen the results are 77% and 72% for correct classification, which is good result for SVM classifiers in general.

4. Overview

Although though better results can be achieved using neural networks the latter require significantly more time to train, and much higher volume of training dataset. Current results can be further improved by adding more features, in example spectral centroid, peaks and spectral bandwidth for different sub bands of the signal. Still hog descriptors the perform stable and should be considered solid foundation of classification features in any classifier. Also differ classifiers could be used after consideration their pros and cons, as well the training dataset volume.

The results are aimed at the implementation of Work Package 2 "Intelligent Security Systems" of project BG05M2OP001-1.002-0006 "Construction and development of the Competence Center "Quantum Communication, Intelligent Security Systems and Risk Management (Quasar)", which has received funding from the European Regional Development Fund through the Operational Program "Science and Education for Smart Growth" 2014-2020.

5 References

1. M. Cowling and R. Sitte. Comparison of techniques for environmentalsound recognition. Pattern recognition letters, vol. 24, no. 15, pp. 2895–2907, 2003.

2. J. Dennis, H. D. Tran, and H. Li.Spectrogram image feature forsound event classification in mismatched conditions. IEEE Signal Process. Lett., Feb. 2011, pp. 130-133

3 N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol.1.IEEE, 2005, pp. 886–893.

4. M. Cowling and R. Sitte.Comparison of techniques for environmental sound recognition. Pattern recognition letters, vol. 24, no. 15, pp. 2895–2907, 2003.

5. D. Wang and G. Brown, Eds., Computational auditory scene analysis: Principles, algorithms and applications. Wiley-Interscience, 2006.

6. G. Yu and J. Slotine. Audio classification from time-frequency texture. in in Proceedings of IEEE International Conference in Acoustics, Speech and Signal Processing, 2009

7. Lichkov N. G, Dimitrov DL, Mihail Hristov, Stoyanov I .. Classification of chain technology as a source of propagation of surface seismic oscillations in different types of soils .. Proceedings of the International Scientific Conference on Security "CONFSEC 2020", 4, 1, NTS in Mechanical Engineering, 2020, ISSN: 2603-2945, 21-24

8. Lichkov N. G, Dimitrov DL, Mihail Hristov, Stoyanov I .. Classification of wheeled vehicles as a source of propagation of surface seismic oscillations in different types of soils .. Proceedings of the International Scientific Conference on Safety "CONFSEC 2020", 4, 1 , NTS in Mechanical Engineering, 2020, ISSN: 2603-2945, 17-20.