PITCH RANGE BASED ENVIRONMENTAL NOISE CLASSIFIER

Burak Uzkent, Buket D. Barkana, Ph.D.

Department of Electrical Engineering, School of Engineering, University of Bridgeport, Bridgeport, CT

SPRG Lab website: http://www1bpt.bridgeport.edu/soe/dsp/index.html

Abstract

Automatic noise recognition was performed in two stages: feature extraction based on the pitch range by analyzing the auto-correlation function and classification using the classifier trained on the extracted features. Since most environmental noise types change their acoustical characteristics in time, we focused on the "pitch range" of the sounds in order to extract features. Two different classifiers, support vector machines (SVMs) and k-Means clustering, were performed and compared using the proposed features. The SVMs and k-means clustering classifiers achieved the recognition rates up to 95.4% and 92.8%, respectively. Although both classifiers provided high accuracy, the SVMs-based classifier outperformed the k-means clustering classifier.

Introduction

Environmental noise recognition has become an important research area. The importance of such research is rising in diverse fields of application such as speech recognition, environmental sound recognition, echo canceller, active noise control, human-machine interface and so forth.

A noise event is defined as an increase of the acoustic level, which is generally unexpected. To identify the source of the noise event is called as Automatic Noise Recognition (ANR). For this purpose, different classifiers based on Artificial Neural Networks (ANN), Hidden Markov Models (HMM), Fuzzy Logic systems, Support Vector Machines (SVMs), and k-Means clustering were developed using traditional or novel feature extractions

Feature selection is one of the most important tasks in a classification algorithm. It allows for a low computational load without increasing the misclassification error. The aim is to obtain an efficient and small vector of acoustic features which represent the input pattern for the classification algorithms being trained.

The classifier developed in this paper is proposed for the recognition of a limited number of classes of environmental sounds –engine, restaurant, and rain- and is motivated by the context aware and audio surveillance applications. Such applications, which are able to classify a number of different environmental sounds related to daily life events, can advance the quality of people's life. We propose a new set of feature parameters based pitch which is different than the traditional feature parameters in previous studies. The performance of the proposed new set of feature parameters is evaluated by the SVMs and the k-Means classifiers.

Feature Extraction

Pitch analysis of audio signals is useful for many purposes. Its applications include automatic music transcription, speech separation, structured audio coding, and music information retrieval. Pitch is a perceptual feature of sounds. Its perception plays an important part in human hearing and understanding of different sounds.

As a part of this study, we calculated feature parameters which characterize different noise types efficiently for classification purposes. Quasi-stationary noise types, rain, restaurant, and engine were used. Using a time-domain technique, the Autocorrelation Function (ACF), pitch is estimated. The equation of the ACF is written below, where is the signal and is the window size. It measures the extent to which a signal correlates with a time offset (τ) version of itself. Because a periodic signal will correlate strongly with itself when offset by the fundamental period, we can expect to find a peak in the ACF at the value corresponding to a period .

$$\phi(\tau) = \sum_{n=o}^{N-1} x(n)x(n+\tau)$$

Two feature parameters are calculated using the pitch range and its mean value for every noise instance:

1) Feature parameter 1 is chosen as the ratio of the maximum and minimum of the mean value (blue-solid line in graphics).
2) Feature parameter 2 is chosen as the ratio of the standard deviation of the pitch range and mean value.

Original signal waveforms and their spectrograms to show diversity among classes

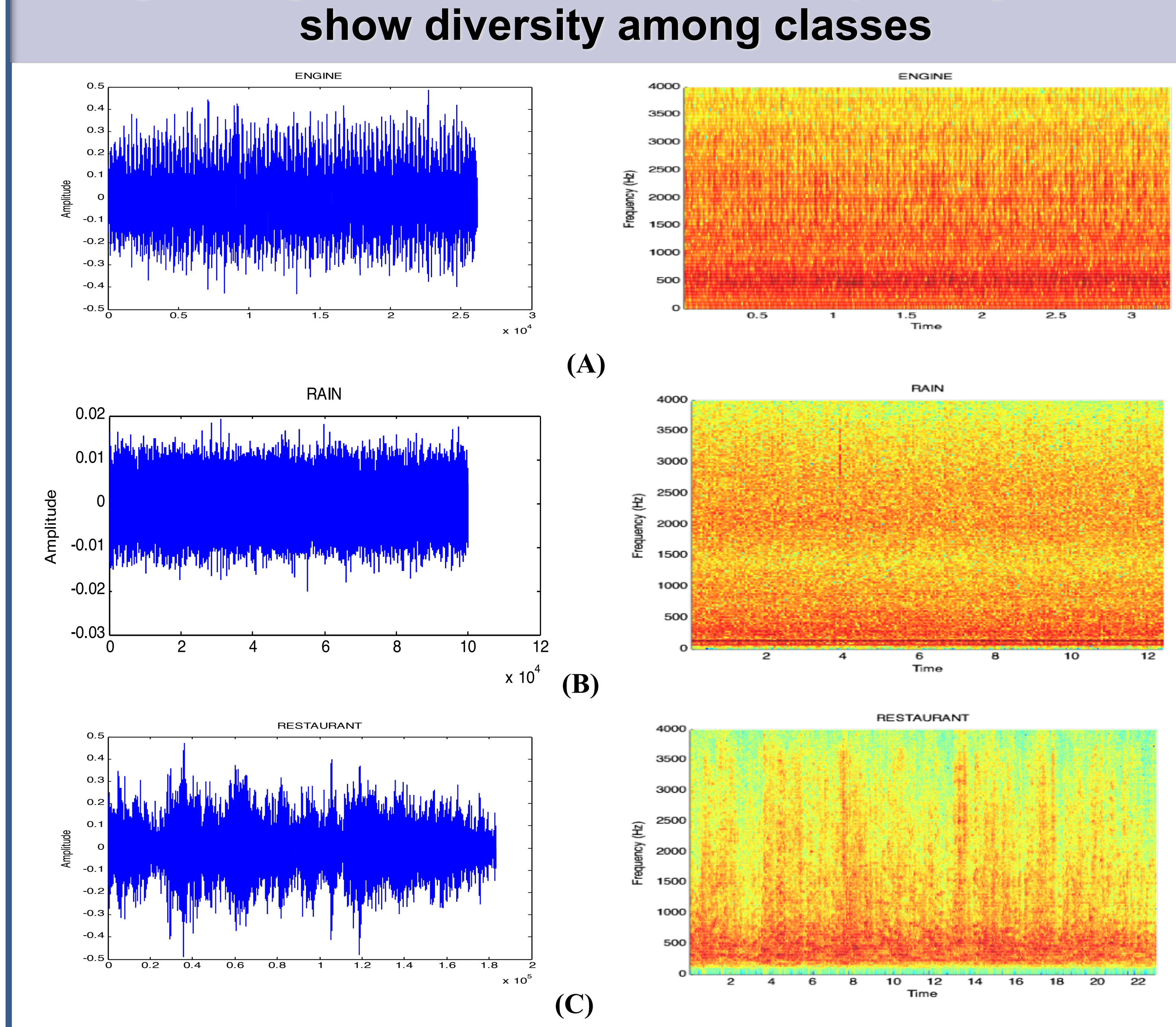

Figure 1: The original waveform and their spectrograms of (A) Engine, (B) Rain, and (C) Restaurant samples

Methods

Database: 19 Engine, 52 Rain, and 28 Restaurant noise instances were analyzed at 96 KHz sampling frequency. All signals in the database have a 16 bits resolution. To calculate pitch, at least 4.69 seconds of each instance was processed using 2.1 msec-rectangular windows with %50 overlapping. Since SVMs requires small amount of training data, %75 of the overall data belongs to testing part. Polynomial and RBF kernel functions were used with three different degrees: Figures 2,3 and 4 show the SVM classifier applied to a rain, restaurant, and engine test set of nonlinearly separable data using the mapping degree as 2. Table 2 shows the classification accuracy.

SVMs: The SVM is based on structural risk minimization (SRM) where the aim is to learn a classifier that minimizes the bound of the expected error . In other words, it seeks a maximal margin separating hyperplane with maximum margin to the closest point of the training set between two classes of data.

K-Means Clustering: It tries to minimize the sum of squares of distances between data and the matching cluster centroid. To get a more accurate output the centroids should be placed far away from each other as much as possible. This process is repeated iteratively until the place of centroids get equal to previous ones.
We used the same database in the SVMs classifier to test the performance of the k-means clustering based classifier. Classification of Rain, Restaurant and Engine data is depicted in Figure 5 the classification accuracy is shown in Table 1.

Results

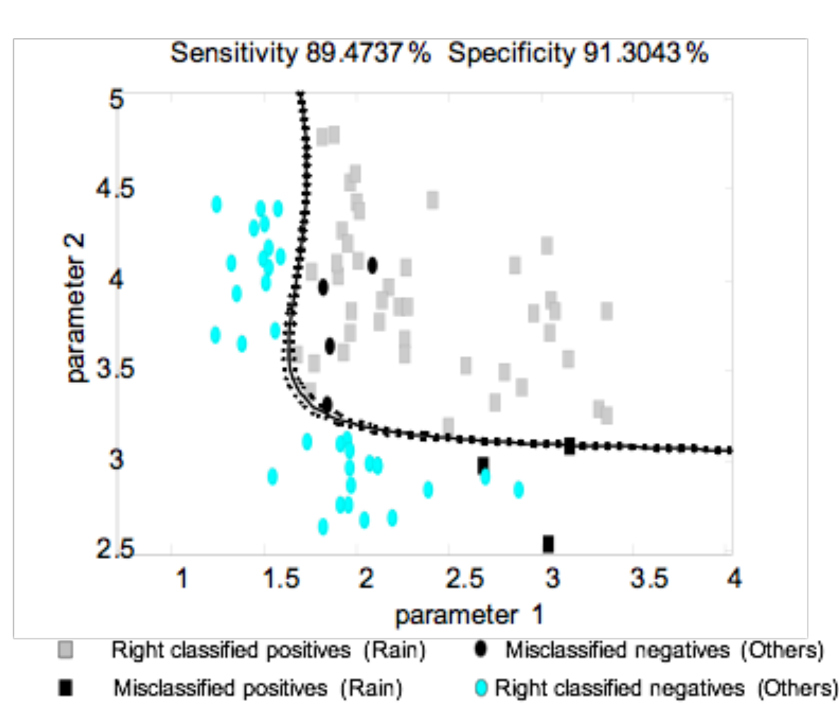

Figure 2: The classification of rain noise type with RBF

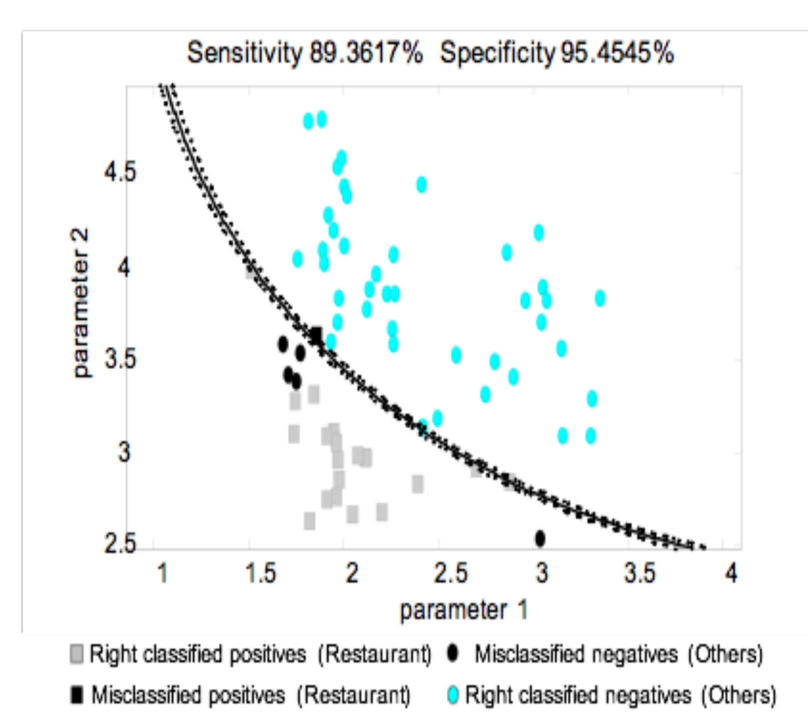

Figure 3: The classification of restaurant noise type with RBF

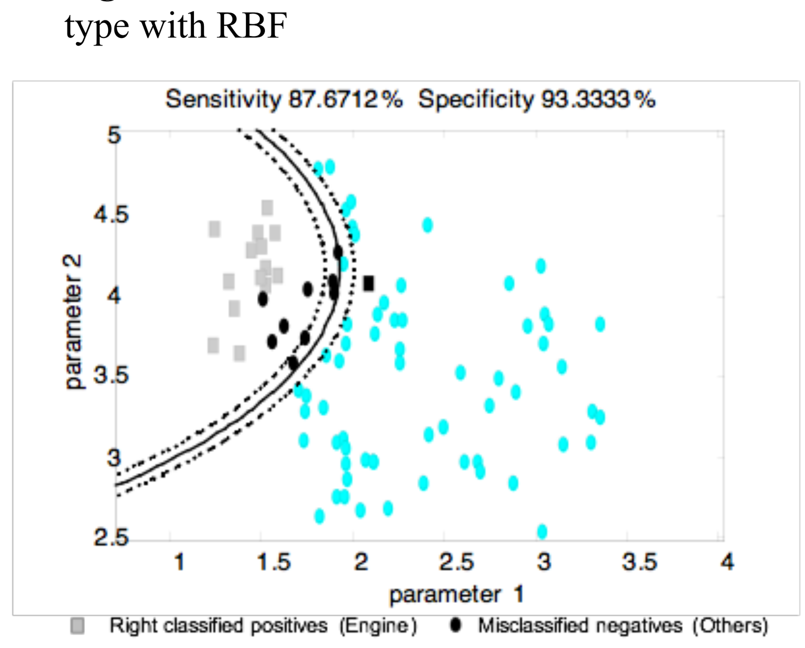

Figure 4: The classification of engine noise type with RBF

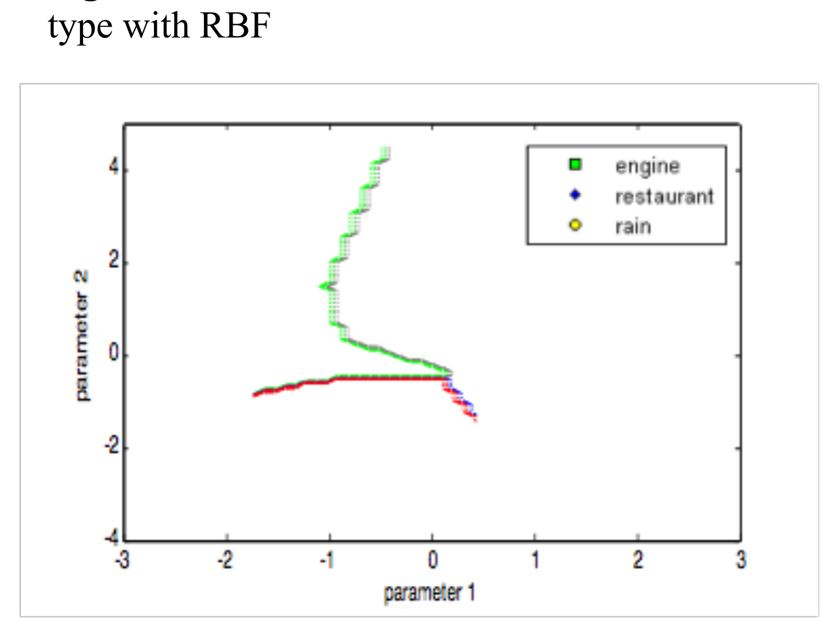

Figure 5: The classification of data (Rain +Engine +Restaurant) with K-means

Conclusion

Automatic environmental noise recognition has been studied with SVMs and k-Means clustering algorithms using newly proposed feature extraction. The extracted feature vectors were based on the pitch calculations. The SVMs classifier showed a great performance since it maps the features to a higher-dimensional space. Its recognition rates ranges between 91.3% and 95.4%. Beside this, the k-Means clustering classifier provided recognition rates between 81% and 92.8%. Overall, the presented classifiers are found effective to improve the performance of the ANR systems

Table 1: Confusion Matrix for K-Means Clustering

True Class	Predicted Class (%)		
	Eng.	*Rain*	*Rest.*
Engine	**92.8**	7.2	0
Rain	9	**81**	10
Restaurant	16	0	**84**

Table 2: Confusion Matrix for SVMs with Radial Basis Function

True Class	Predicted Class (%)								
	d=1			d=2			d=3		
	Eng..	*Rain*	*Rest.*	Eng	*Rain*	*Rest.*	Eng.	*Rain*	*Rest.*
Engine	**93.3**	6.67	0	**93.3**	6.67	0	**93.3**	6.67	0
Rain	2	**93.5**	4.5	0	**91.3**	8.7	0	**93.4**	6.53
Restaurant	9.1	0	**90.9**	0	4.6	**95.4**	0	9	**91**