Identification of acoustic signatures for vehicles via reduction of dimensionality

Amir Averbuch[1] Eyal Hulta[1] Valery Zheludev[1]
[1]School of Computer Science
Tel Aviv University, Tel Aviv 69978, Israel

April 15, 2001

Abstract

In this paper we propose a robust algorithm that solves two related problems: 1) Classification of acoustic signals emitted by different moving vehicles. The recorded signals have to be identified to which pre-existing group they belong to independently from the recording surrounding conditions. 2) Detection of the presence of a vehicle in a certain class via analysis of its acoustic signature against the existing database of recorded and processed acoustic signals. To achieve this detection with no false alarms we construct the acoustic signature of a certain vehicle using the distribution of the energies among blocks which consist of coefficients of multiscale local cosine transform (LCT) applied on the frequency domain of the acoustic signal. No false alarms in the detection are allowed even under severe conditions such as the acoustic recording of the sought-after object is a superposition of the acoustics emitted from other vehicles that belong to other classes and external noise. The proposed algorithm is robust even under severe noise and diverse rough surrounding conditions. This technology, which has many algorithmic variations, can be used to solve wide range of classification and detection problems which are based on unique derivation of signature.

1 Introduction

In the paper we address the problem of detection of the presence of a specific vehicle from a certain class of vehicles via analysis of its acoustic signature against an existing database of related vehicles. This problem and its more general setting for classification of acoustic signals emitted by different moving vehicles are complex because of the great variability in the surrounding conditions of the recorded database such as velocity, distances between the vehicles and the receiver, the presence of other vehicles, the roads the vehicles are traveling on, the background noise, just to name few.

A crucial factor in having a successful classification (no false alarms) is to construct unique and reliable signatures that contain the right unique characteristic features that enable to discriminate among recorded classes. Our approach relies on the fact that each part of the vehicle emits a distinct acoustic signal which in frequency domain contains only a few dominating bands. The basic assumption is that the acoustic signature for the class of signals emitted by a certain vehicle is obtained as a combination of the inherent energies in a small set of the most discriminant frequency bands. As the car moves the conditions are changed and the configuration of these bands may vary, but the general disposition remains. To reveal these characteristic frequency bands we first apply the discrete Fourier transform on the signals and then use the multiscale local cosine transform (LCT) [9, 17, 18, 19] on the frequency domain of the signal. The blocks of the LCT coefficients, each of which is related to a certain frequency band, are the relevant tool to base the classification on. These blocks contain the distinctive characteristic features. To select the blocks of the LCT coefficients which most discriminate the class of interest from other classes, we use methods similar to the Local Discriminant Basis (LDB) algorithm by Saito and Coifman [21, 22, 23].

Another way to extract these characteristic features is to apply the wavelet packet transform on the time domain of the signal and to use as characteristic features energies in the blocks of wavelet packet coefficients [1]. As a matter of fact, this paper solves differently the same problem as stated in [1].

In the final phase of the process, in order to identify the acoustic signatures of the predetermined classes of signals, we compare between two classifiers: 1. The conventional classifier that is based on Linear Discriminant Analysis (LDA) [8]. 2. A classifier which is based on "parallel coordinates" concept devised by Inselberg [11, 12, 13] which is a visualization tool that automatically enables to reduce the dimensionality of the problem. In all our experiments the parallel coordinates classifier overwhelmingly outperformed the LDA.

The algorithms were tested on a wide series of field experiments. The results demonstrate that certainly the proposed algorithm is robust and the false alarms rate is near zero.

The paper is organized as follows.

2 Formulation of the approach

2.1 The structure of the signals

The raw data that was used to train the algorithm is a collection of acoustic recordings of four vehicles. The goal of this paper is to have a unique identification of the vehicle labeled G^1. As a reference set we used recordings that were taken from three different vehicles labeled $G^{2,3,4}$. The signal from each vehicle was recorded while traveling at typical driving velocities. The data was collected in field conditions, while the different vehicles traveled toward and away from the recording unit up to a distance of 1500m. The acoustic signal was processed by a linear phase analog filter (cutoff at 512Hz) and then digitized at 1KHz with a 12 bit digital recorder. This cutoff frequency was selected since the unique acoustic tones emitted by the vehicle are all within the range 0 to 512Hz.

The basic time fragment for the analysis is called *slice* and labeled by S. Each slice is 0.25 seconds long ($2^8 = 256$ samples). The time interval is selected to be 0.25 of a second since the acoustic signal is considered to be quasi-periodic. During such a time interval the spectrum of the signal varies insignificantly. The Fourier transform of a slice $S(t)$, denoted by $\hat{S}(\omega)$, contains 128 bins each 4Hz width. Figure 1 presents time slices for each of the four vehicles. Figure 2 presents the spectra of these slices.

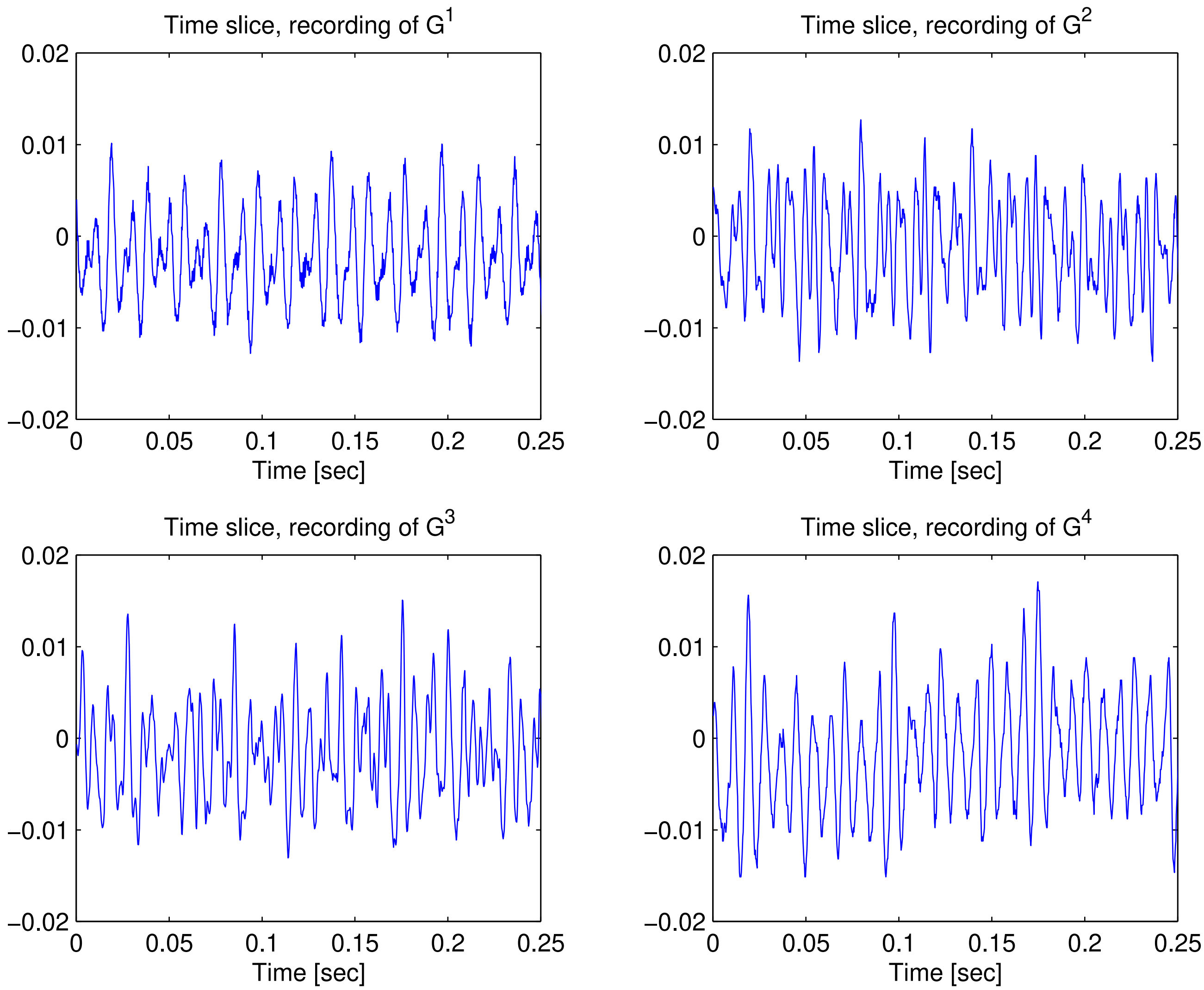

Figure 1: Time slices from the four vehicles

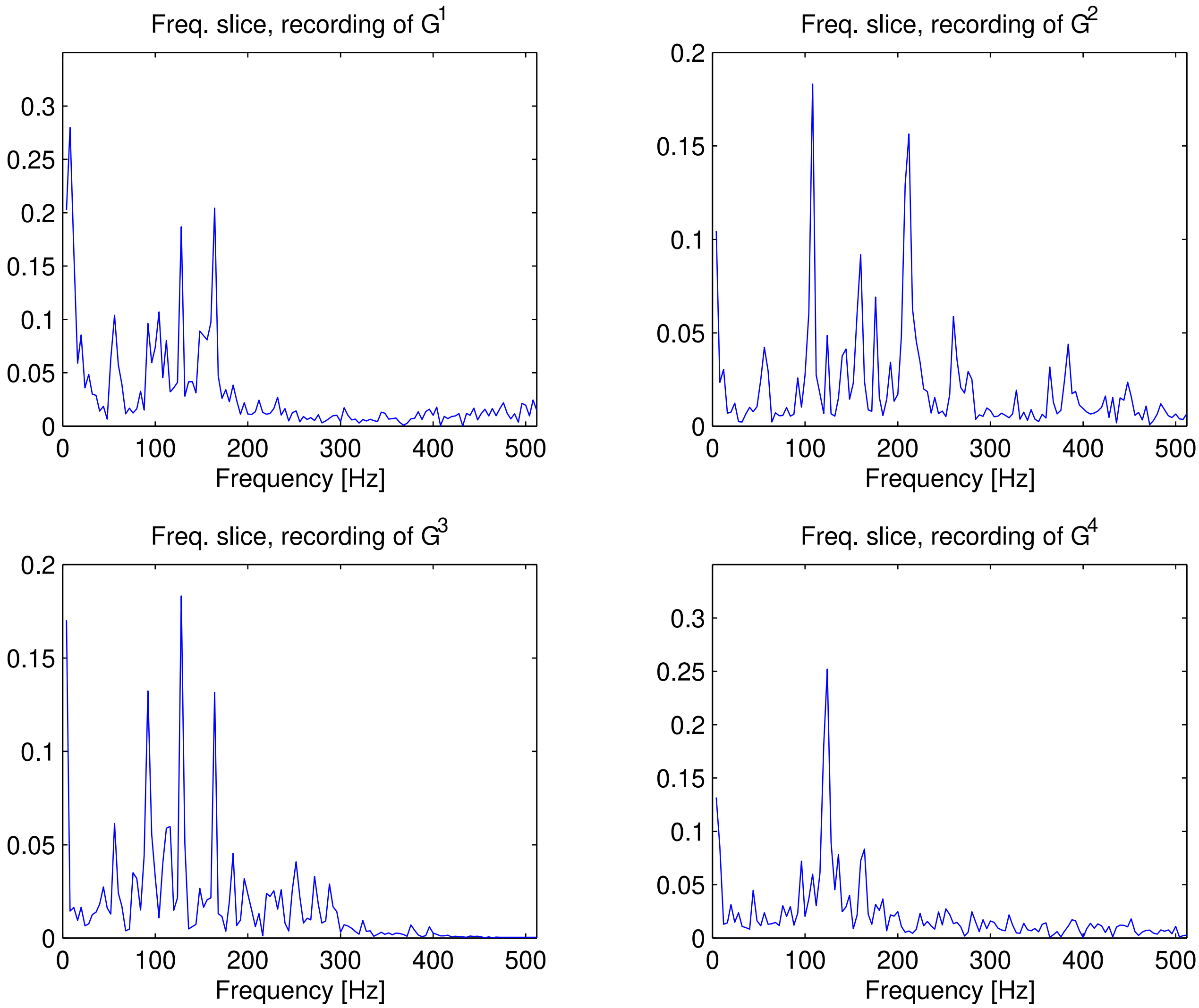

Figure 2: Frequency domains of slices from the four vehicles

We can see from Fig. 2 that each of the vehicles possesses a unique set of acoustic frequencies (or frequency events). As mentioned before, each of those frequency events is an acoustic feature. Each feature deviates in time around some center frequency and within a frequency band.

In order to better visualize the features, Fig. 3 presents an histogram of the frequency domain based on the recordings of the vehicle G^1. The histogram is constructed by Fourier transforms of consecutive slices. Each slice is 0.25 second in time, and consecutive slices overlap by 75% (hence we have 16 slices per second). The result is a snapshot of the frequency domain as function of time. The basic acoustic features of the sound emitted by a vehicle are clearly visible (the sliding frequency ridges).

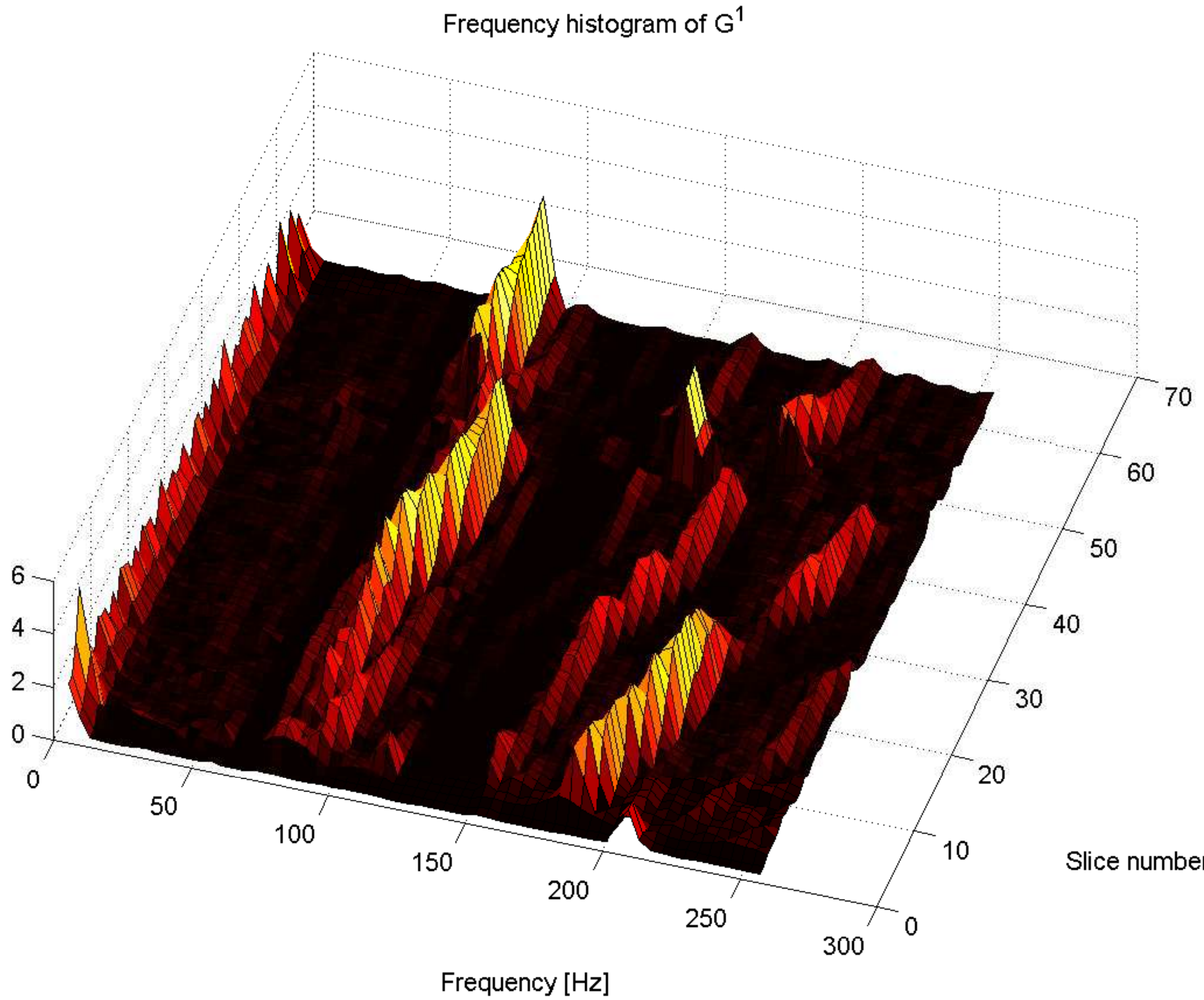

Figure 3: Time-frequency histogram of vehicle G^1. Each row in the figure is a frequency domain of one slice recording. Each slice is 0.25 second in time and the presented frequency domain is 256Hz wide (64 bins, 4Hz per bin). The sliding ridges are the acoustic features (tonal lines) whose instantaneous frequency changes within the frequency band.

2.2 Formulation of the approach

The problem which we are solving can be formulated as a classification into two classes:

Class C^1 which contains signals emitted by the G^1-vehicles at various velocities, distances and types of roads.

Class C^2 which contains signals emitted by other different types of vehicles.

Due to the quasi-periodic structure of the signals, local cosine transform (LCT) [9] analysis applied on frequency domain of the signals is a relevant tool for adaptive search for valuable frequency bands of a signal or a class of signals. Once implemented, the local cosine transform yields a huge variety of different versions for the partition of the frequency domain. One of the important features of the transform is its computational efficiency. Following the suggestion in [15] we increase the number of sample in the training sets and in the test sets by imposing a comparatively short window on each input signal followed by a shift of this window along the signal so that adjacent sections have some overlap.

The one-dimensional LCT is applied in a multiscale fashion to perform adaptive segmentation of the input signal. As explained in [5] we can adaptively select the size and location of the windows with the best basis algorithm. We consider a function with finite support. We divide the support into two intervals, and we consider the local cosine basis associated with each interval. We then further decompose each interval into two sub-intervals, and consider the local cosine bases associated with this finer subdivision.

By applying this decomposition recursively we obtain a homogeneous binary tree-structured decomposition as shown in figure 4.

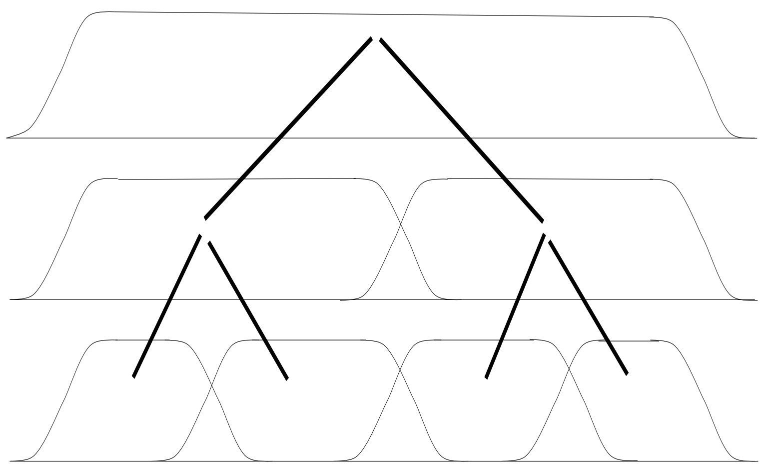

Figure 4: Binary tree decomposition. Within each interval the signal is expanded into a local DCT. The optimal segmentation is then searched for.

For each interval, or node in the binary tree, we calculate the set of coefficients in the subblock. If we associate a cost for each node of the tree, based on the set of coefficients, then we can find an optimal segmentation. Using a divide and conquer algorithm, group of connected nodes are pruned if their total cost is greater than the cost of their father. The process is recursively applied from the bottom to the top of the tree, and a global optimal tree is then found.

In our computation of the acoustic signals L_2 based cost function was chosen.

The general approach:

1. For training we use a set of Fourier transformed signals with known membership. From them we select a few blocks with multiscale LCT coefficients which discriminate efficiently between the signals emitted by the vehicle G^1 and signals from other vehicles.

2. We apply the Fourier transform and multiscale LCT on the signal to be classified. We use as its characteristic features the normalized l_2 norms of the LCT coefficients

contained in the selected blocks.

3. Finally, we submit the vectors of the extracted features to the "Parallel Coordinates" classifier. It was trained beforehand. It decides which class this signal belongs to.

3 Description of the algorithm and its implementation

3.1 The algorithm

The algorithm is centered around two basic operations:

I. Selection of the discriminant blocks of the multiscale LCT coefficients. This is performed according to the following steps:

1. Construction of the training set.
2. Calculation of the energy map.
3. Evaluation of the discriminant power of the decomposition blocks.
4. Selection of the discriminant blocks.

II. Discrimination among the signals.

1. Preparation of the pattern set.
2. Building the classifier based on the "Parallel Coordinates" methodology.
3. Preparation of the testing set.
4. Making the decision.

Now we present a detailed description for the implementation of the algorithm.

3.2 Implementation of the algorithm

3.2.1 Selection of discriminant blocks

The technique we use for the selection of the discriminant blocks is similar to the Local Discriminant Basis algorithm by Saito and Coifman [21, 22, 23].

Construction of the training set: Initially, we collect as many as possible recordings of the signals emitted by vehicle G^1 (Class C^1signals) and signals from vehicles G^l, $l = 2, 3, 4$ (Class C^2 signals). We prepare from each selected recording, which belongs to a certain class, a number of overlapping slices of length 256 samples each, shifted with respect to each other by 64 samples. All together we produce M^l slices $\{S_i^l\}_{i=1}^{M^l}$ for the class C^l, $l = 1, 2$. These groups of slices form the training set for the search of the discriminant blocks.

Calculation of the energy map: It is carried out in the following way:

1. The power spectrum of each slice from a given class C^l is calculated. We obtain M^l "frequency slices" $\{F_i^l\}_{i=1}^{M^l}$ (bins) of length $n = 128$ for the class C^l. This operation is followed by application of the multiscale local cosine transform on slices $\{F_i^l\}$ up to level (scale) $m < 7$. This procedure produces mn coefficients arranged into $2^{m+1} - 1$ blocks $B_i^{l,s} = \{b_i^{l,s}(j)\}_{j=1}^{R_s}$ associated with different frequency bands for each slice $\{F_i^l\}$.
2. The energy map of the block $B^{l,s}$ is defined as the normalized mean value of its coefficients over all the slices of class C^l:

$$\{E^{l,s}(j)\}_{j=1}^{R_s} = \frac{\sum_{i=1}^{M^l} b_i^{l,s}(j)^2}{\sum_{i=1}^{M^l} |F_i^l|^2}. \tag{3.1}$$

   Similar operations are performed on all blocks of the LCT coefficients of both classes C^l, , $l = 1, 2$.

Evaluation of the discriminant power of decomposition blocks: Let $\{E^{1,s}(j)\}_{j=1}^{R_s}$ and $\{E^{2,s}(j)\}_{j=1}^{R_s}$ be the energy maps of the block s, denoted by B^s, in the classes C^1, C^2. The discriminant power of this block is defined as the sum:

$$D^s = \sum_{j=1}^{R_s} (E^{1,s}(j) - E^{2,s}(j))^2. \tag{3.2}$$

Selection of discriminating blocks: Now we are in a position to select a few discriminant blocks which form the acoustic signatures for the classes. Since we are in a situation where the frequency bands of the blocks overlap and we get overcomplete coverage of the space, we use the *Best Basis Selection Algorithm* [17] to discover in the multiscale expansion which block has a stronger discriminant power. The idea is to compare among the discriminant power of each pair of the "children" with those of their parent in the multiscale expansion. In the case when the discriminant power of the parent exceeds the sum of the children, the children blocks are discarded and vice versa. As a result we obtain some non-overlapping set of blocks that covers the whole frequency domain of our recorded signals. This set is referred to as the "most discriminating basis". Typically, this set contains relatively large number of blocks, especially, if the depth m of the decomposition is large. Therefore, we select from this set a few blocks B^s, $s = 1 \dots, M$ with the highest discriminant factor. Moreover, if we are interested in certain frequency bands, we can select the corresponding blocks.

As a result of the above operations we discover a relatively small set of LCT blocks with highest the discriminant power. This part of the investigation is computationally expensive, especially if, for better robustness, large training sets are involved. On the other hand this task is performed once as a preprocessing phase.

3.3 Classification

Once we have the set of discriminant blocks B^s, $s = 1 \dots, M$, we proceed to the classification phase.

3.3.1 Preparation of the reference set

Initially, we choose a number of recordings that belong to the classes C^l, $l = 1, 2$, to be distinguished, from which we form the reference set. These recordings are sliced similarly to those which were used for the preparation of the training set. For a certain class C^l, we form from all the selected recordings from the class a number of overlapping slices each of length $n = 256$ which corresponds to 1 second. These slices are shifted with respect to each other by 64 samples. We assume that there are μ^l slices $a^l(i,:), i = 1, \ldots, \mu^l$, related to the class C^l. We calculate the power spectrum of each slice and gather the obtained "frequency slices" into $\mu^l \times n/2$ matrix F^l, $l = 1, 2$.

Then, we apply the multiscale LCT up to scale m on each row $F^l(i,:)$ of this matrix. After the decomposition of the frequency slice $F^l(i,:)$, we calculate the energies of the M blocks B^s, $s = 1 \ldots, M$ that were selected before. The energy of a block $B^s = \{b^s(j)\}_{j=1}^{R_s}$ is defined as the sum of squared magnitudes of its terms:

$$E^s = \sum_{j=1}^{R_s} b^s(j)^2. \tag{3.3}$$

In doing so we obtain the $1 \times M$ vector $\vec{v}^l(i) = \{v^l(i,j)\}$, $j = 1, \ldots, M$. In order to place all the vectors into a M-dimensional hyperplane, we apply the following normalization:

$$\vec{V}^l(i) = \{V^l(i,j)\}_{j=1}^{M} = \left\{ \frac{v^l(i,j)}{\sum_{k=1}^{M} v^l(i,k)} \right\}_{j=1}^{M}.$$

We regard the vector $\vec{V}^l(i)$ as a representative of the slice $a^l(i,:)$. The vectors $\vec{V}^l(i)$ form the $\mu^l \times M$ reference matrix V^l of the class C^l. We do the same for both classes C^l, $l = 1, 2$.

The reference sets are used for two purposes: **1.** As pattern sets for the LDA classifier. **2.** As training sets for building the classifier which is based on the *Parallel Coordinates* method.

3.3.2 Classifier construction - extraction of the acoustic signature

The goal now is to extract the acoustic signature of the class C^1. The representative vectors for slices originated from class C^1 are embedded in an M-dimensional hyperplane space. We have to find a domain in this hyperplane that contains as many vectors from class C^1 as possible, and as few vectors as possible from the other class C^2. This domain is reffered to as the *acoustic signature* of class C^1.

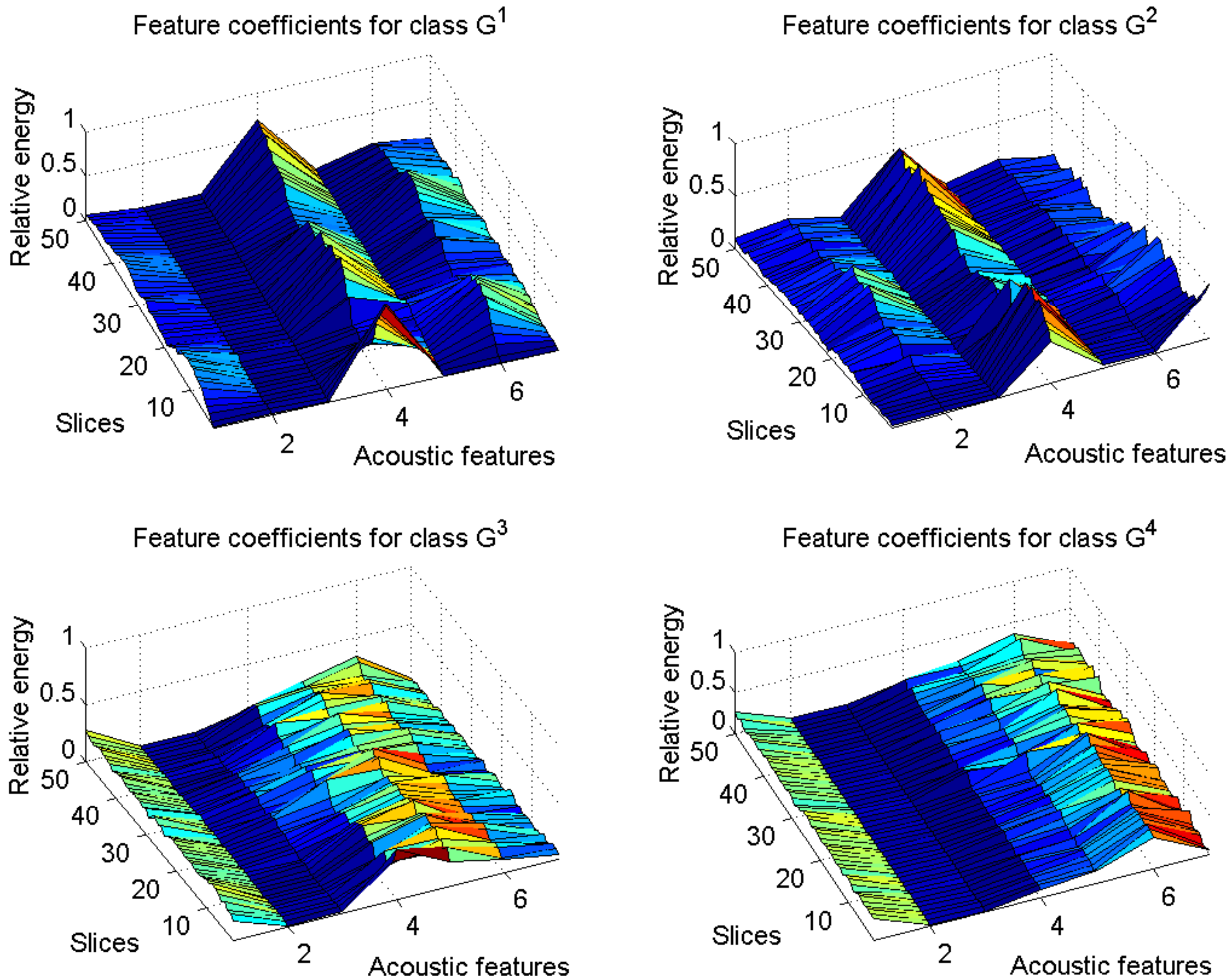

Figure 5: Acoustic feature coefficients of classes $G^{1,2,3,4}$ vehicles. For each vehicles consecutive slices were decomposed using LCT and the 7 blocks were extracted and normalized. The coefficients of each slice are presented by a polygonal line, and a histogram of those values over time is built for each vehicle.

In this section, we present our method for identifying the acoustic signature. We do so by training a classifier with data taken from the vehicles $G^{1,2,3,4}$. Efficient visualization tool for a multidimensional space can help in identifying the smallest necessary feature space that contains the signature. The need for a visualization tool *that reduces the dimensionality of the problem* is clear when the limits of ordinary classifiers, such as the LDA classifier, are fully understood. Our method for extraction of the acoustic signature is based on the parallel coordinates analysis which is implemented by the "Parallax" software [11, 12, 13]. This analysis offers a systematic, automatic and robust way to visualize the M-dimensional feature space. In short, the characteristic domain of each class has to be *seen* and then "framed" in $M - dimensional$ "boxes". By using the *point-line dualism*, each vector in the M-dimensional feature space is plotted as a polygonal line between M parallel coordinates.

Four pictures in Figure 6 present in parallel coordinates 7-dimensional subspaces containing the signatures of vehicles of four types.

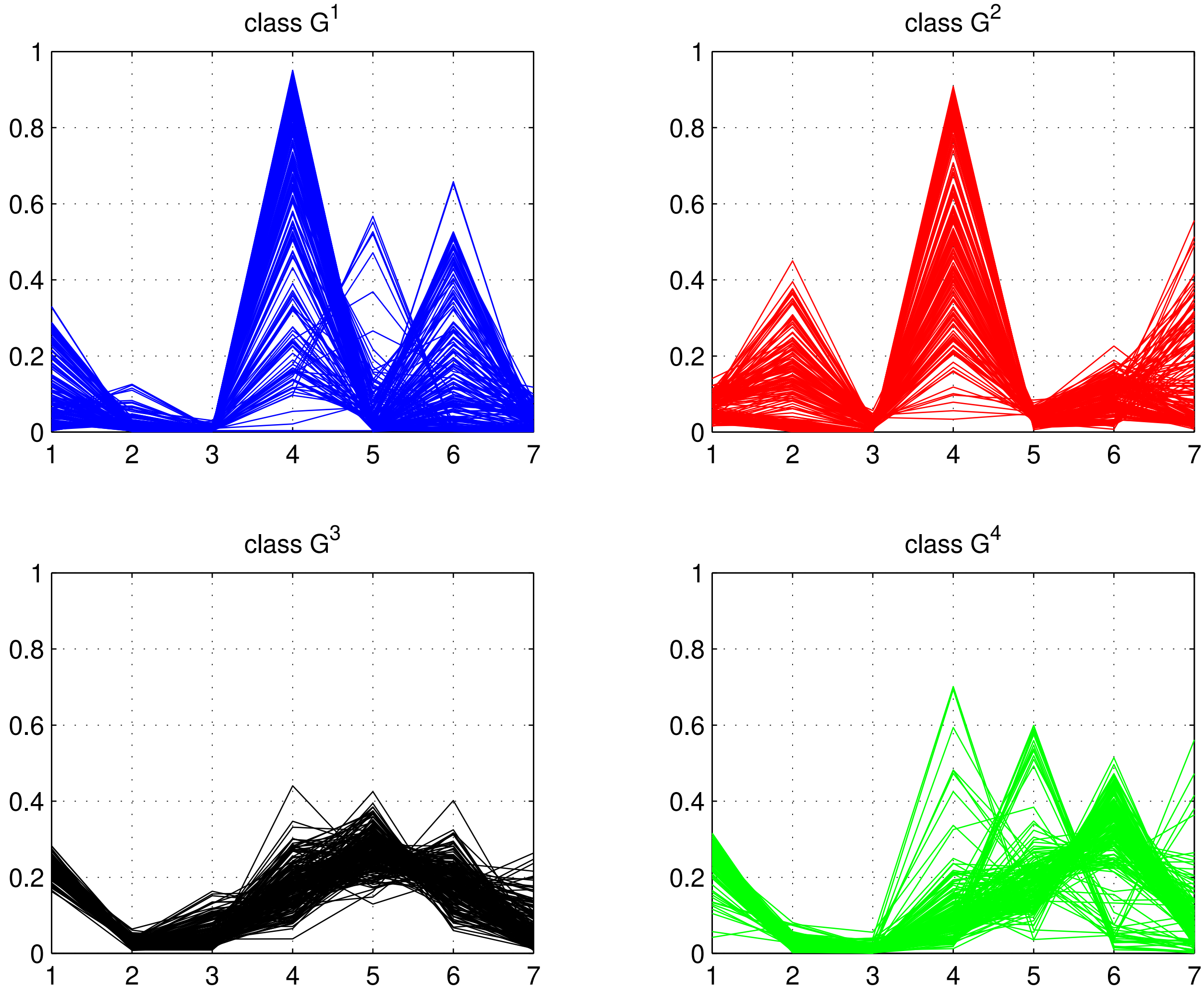

Figure 6: Parallel coordinate plot in 7-D of the vectors $G^{1,2,3,4}$. Each class is presented in a different subplot. Each slice is transformed into a vector in 7-D (seven block coefficients). Such a vector is drawen as a polygonal line using seven parallel coordinates.

The problem then is reduced to framing a subset in the hyper-plane, that contain vectors representing slices of the sought after vehicle and does not contain vectors related of other vehicles. Such a subset of a hyper-plane is confined by a lower and upper bounds in each coordinate, resulting in an M-dimensional *box*. The boxes that cover best the unique volume of the vehicle comprise its acoustic signature. The parallel coordinates help to visualize exactly which are the coordinates of the box that easily separate between classes and which coordinates have overlapping values for different classes. The acoustic signature box is first defined by setting the proper bounds to the separating coordinates. If there is no easy way to separate classes then bounds are set so that much of the wanted domain is contained in the box and as much of unwanted domain is excluded. Figures 7, 8, 9 demonstrate the method (using Parallax [13]).

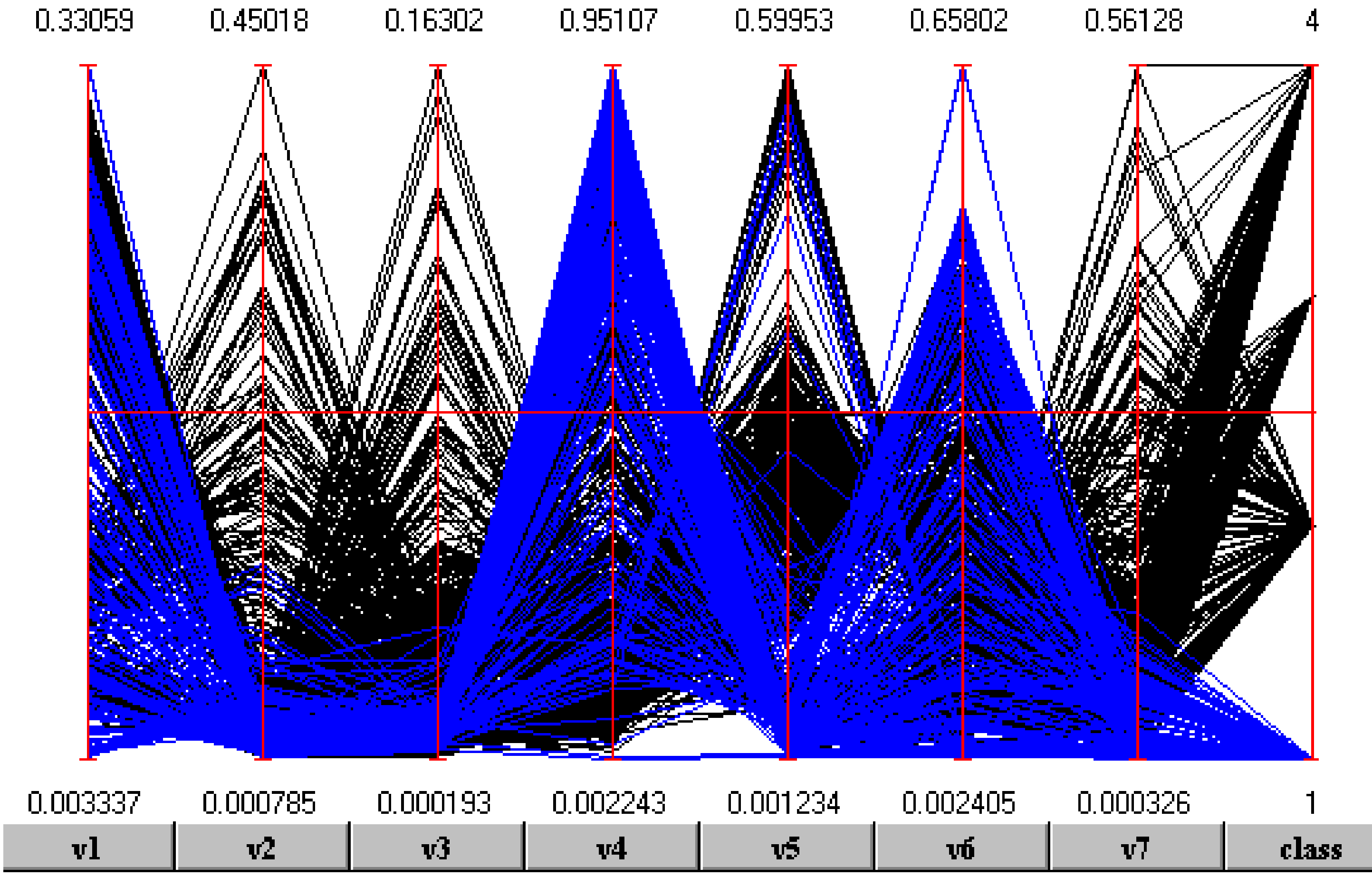

Figure 7: Parallel coordinate plot in 7-D of the vectors of class C^1 (blue lines). The vectors of the other slices are in black. The 7 coordinates are represented by the first 7 columns. The final column (labeled *class*) separates the 4 types of vehicles ($G^{1,2,3,4}$).

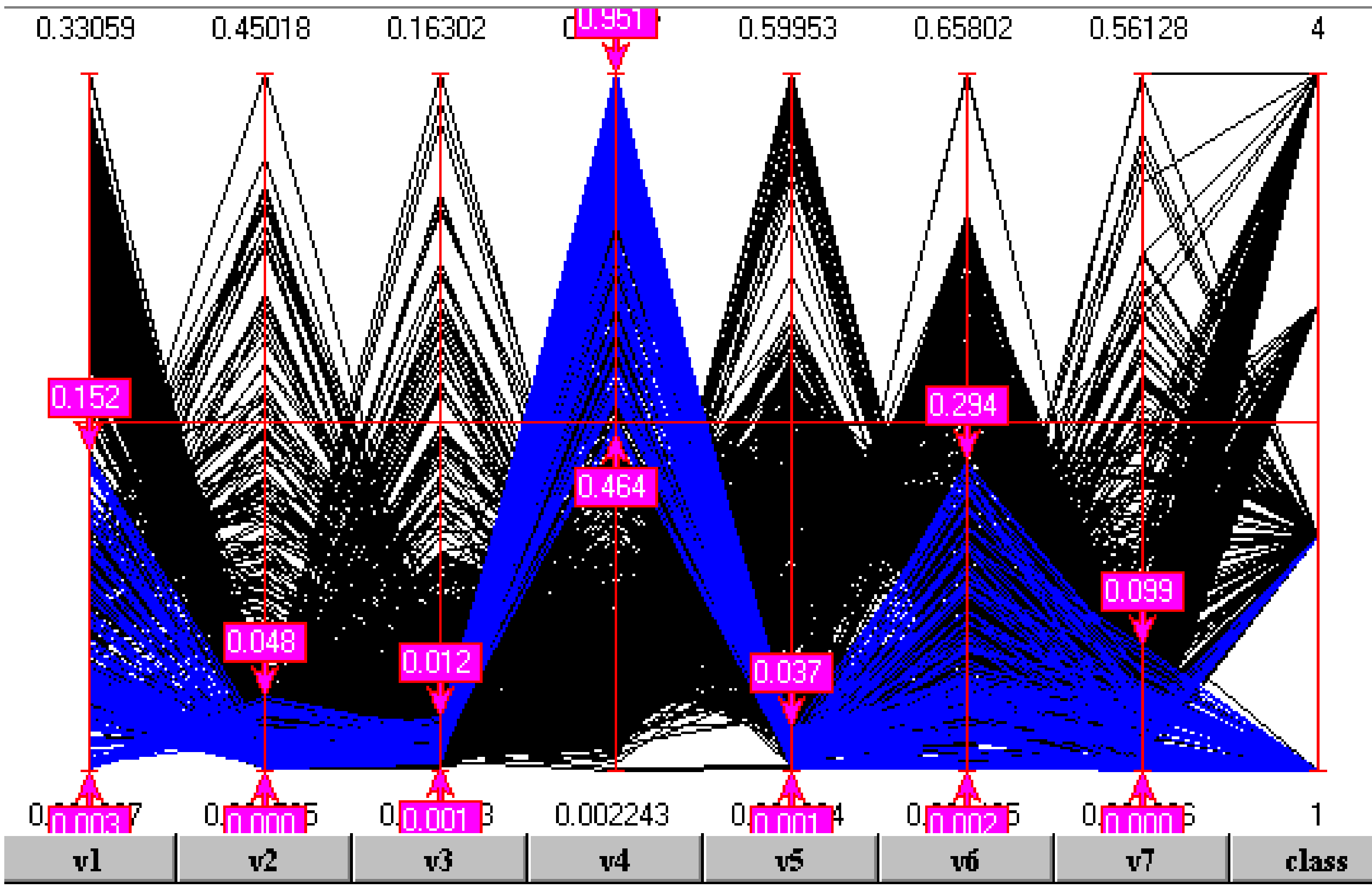

Figure 8: A box framing class C^1. The blue lines represent vectors that are contained in the box, the other vectors are colored black. The box (the bounds in each coordinates are marked by arrows) contains 60% of the vectors of class C^1, and a few vectors of type G^2. The little overlap with G^2 does not cause a false alarm (see figure 13 in section 4).

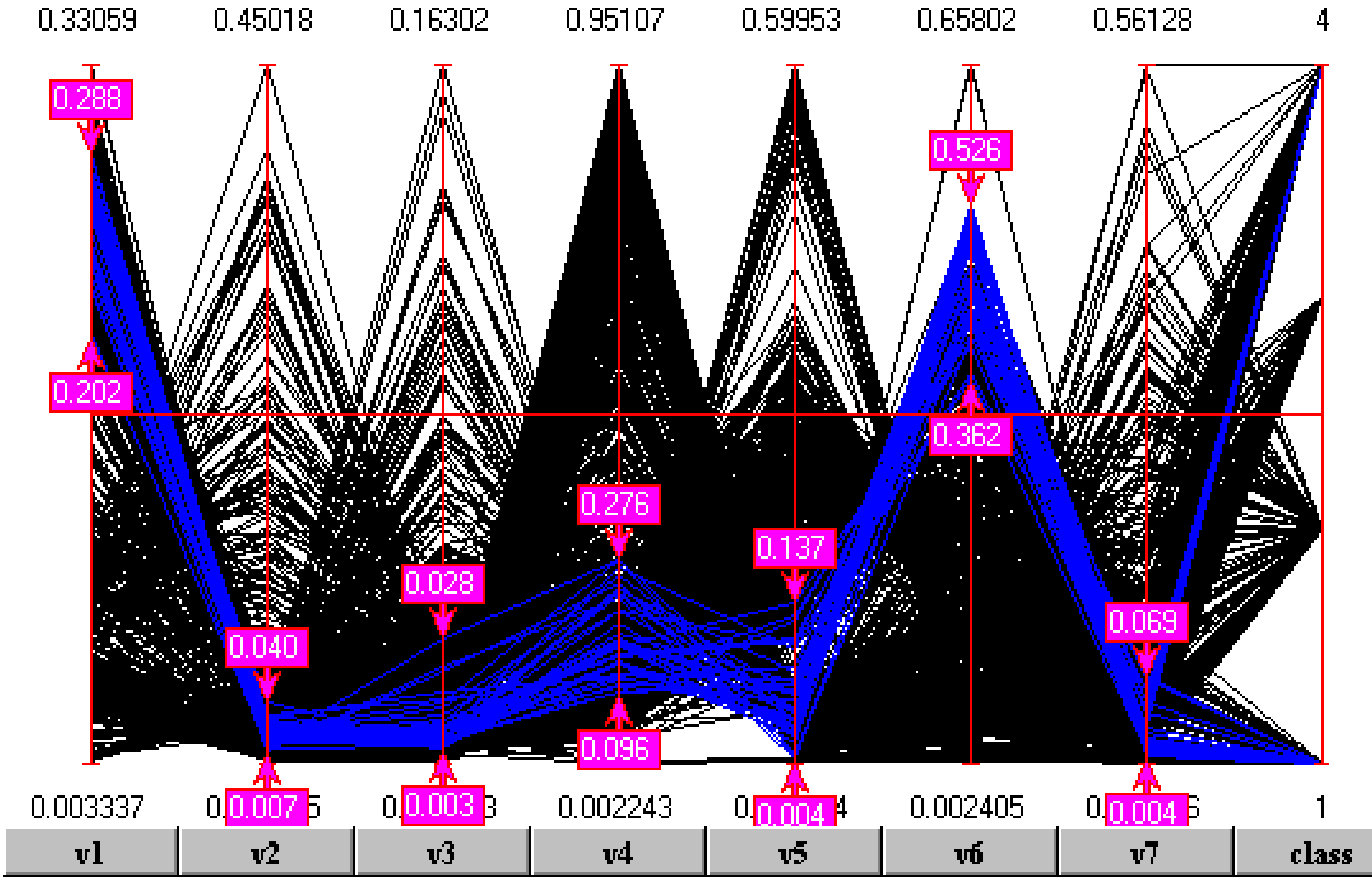

Figure 9: Another box framing class C^1. The blue lines represent vectors that are contained in the box, the other vectors are colored black. The box (the bounds in each coordinates are marked by arrows) contains 20% of the vectors of class G^1 (that are not contained in the previous box), and a very few vectors from class G^4.

This method produces a number of boxes that contain most of the vectors of class G^1. The acoustic signature is then formed as the union of these boxes. In some cases the overlapping between classes makes it difficult to separate between them with boxes. In such a case, we can use a *"hole"*. A hole is a framed box that contains the vectors of the false alarm and that is excluded from the union of boxes. In this case, if a vector is contained in either the union of boxes and the the hole, it is discarded. Finally, the signature is constructed as a union of boxes excluding a union of holes, and a tested vector is either contained or not contained in that domain. The main advantages of this method are:

1. The treatment of false identifications is simple. A vehicle will be falsely identified when the corresponding set of vectors overlaps with that of the sought vehicle. Using the parallel-coordinates, it is easy to isolate that overlapping domain and remove it from the union of boxes using a hole. On the other hand, in the LDA based classification there is no easy way to define a transformation matrix that emphasizes the need to discriminate one of the sets over the others.

2. Unlike the LDA classifier, the parallel-coordinates classifier is not a *relative* classifier since the signature of a class is bounded and framed. In other words, a vector is not classified according to the class it is relatively close to. It is classified as C^1 only if it is contained in a union of framed boxes. As a consequence, an unknown vehicle (that wasn't present in the training phase) is more likely to be classified correctly than while using LDA. Even in the case where an unknown vehicles is falsely identified, it can be removed by the use of holes without replacing the rest of the boxes. Hence, the process of adding a new vehicle does not mean the creation of a totally different signature (unlike the LDA, where the transformation matrix A' needs to be re-calculated according to the mean and variance of the new vehicle's vectors).

On the other hand, the method of defining the boxes and holes is not totally automated as calculating the LDA transformation matrix is. Each of the boxes and holes are found manually through "data mining" in the training set.

After the completion of the construction of the parallel-coordinates based classifier and having pattern sets for LDA, we are in a position to classify test signals. To do so, we must preprocess these signals.

3.3.3 Preparation of the test set

Assume we are given a signal S whose membership in a certain class has to be established. The algorithm has to be capable to process either a fragment of the recording, or the entire recording, or even a number of recordings. We form from the signal S a number K of overlapping slices of length $n = 256$ each (that corresponds to 0.25 second), shifted with respect to each other by 64 samples. So, 16 slices correspond to 1 second recording . After calculation of power spectrum of each of the K slices the results are gathered into $K \times n/2$ matrix T.

The multiscale local cosine transform is operated on each row $T(i,:)$ of this matrix up to scale m. In the decomposed slice $T(i,:)$, we calculate the "energies" of the M discriminant blocks B^s, $s = 1\ldots, M$ that were selected before. In doing so we obtain the $1 \times M$ vector $W(i,:)$ which we regard as a representative of the slice $T(i,:)$. The vectors $W(i,:)$ form the $K \times M$ test matrix W associated with the signal S.

Since the goal of the research is the acoustic signature of a class C^1 vehicle, the recordings are only checked for resemblance to the class C^1 acoustic signature.

3.3.4 Making the decision

Once the test matrix W is ready, we present each row $W(i,:)$ of the matrix to two classifiers:

1. LDA calculates the Mahalanobis distances of the vector $W(i,:)$ from L pattern sets associated with the classes C^l, $l = 1, 2$ and attributes it to the class whose distance was the least.

2. The parallel coordinate classifier is checking whether the vector $W(i,:)$ belongs to the union of boxes which determines the signature of class C^1.

If the vector is attributed to class C^1 then it is assigned the value 1, otherwise – the value is 0. In order to clearly present the changes in identification over time, the original data (16 binary points per second) is processed by a moving average of length 16. In other words, every point on the graph represents an identification result averaged over 16 neighboring slices. This averages the abrupt changes of identification, and the larger the domain below the curve, the better the identification. Hence, when this method is applied on a tested recording, it gives the impression as to how good the identification is Figure 10 demonstrates how this is done.

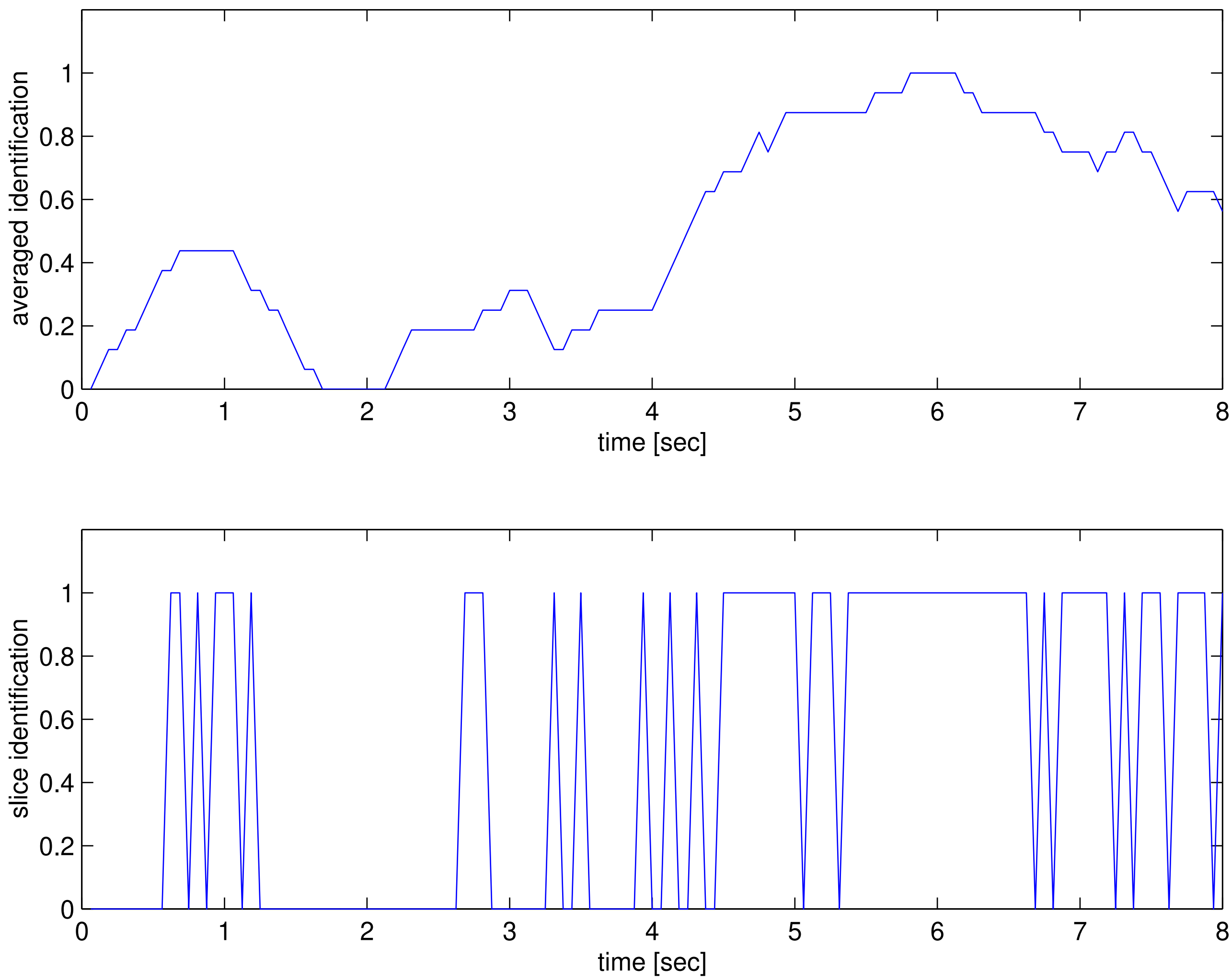

Figure 10: An example of the identification process of a recording. The bottom figure shows whether each vector is found to be a member of the acoustic signature of class C^1. The top figure shows the output of a moving average of the bottom figure..

4 The identification results

We conducted two series of experiments:

1. Closed set identification, in which the 4 types of vehicles ($G^{1,2,3,4}$) participate, and we compare between the LDA and the parallel-coordinates performances. It will be demonstrated that the parallel-coordinate classifier is superior in the separation of the classes.

2. Open set identification, in which a series of known and unknown vehicles will be presented to the parallel-coordinates classifier in order to test its robustness.

In order to clarify the acoustic behavior of a recorded vehicle, a RMS[1] sketch of the acoustic signal of the moving vehicle is presented synchronized with the identification results. Note that normally RMS attains its maximum when the vehicle passes by the recording unit. As the the vehicle recedes from the recording unit, RMS fades.

4.1 Closed set identification

In figures 11, 12, 13 and 14, the identification results from various recordings using the LDA and the parallel-coordinates classifiers are presented.

[1]RMS - root mean square. The root mean square is calculated for each slice and represents the acoustic intensity of the slice

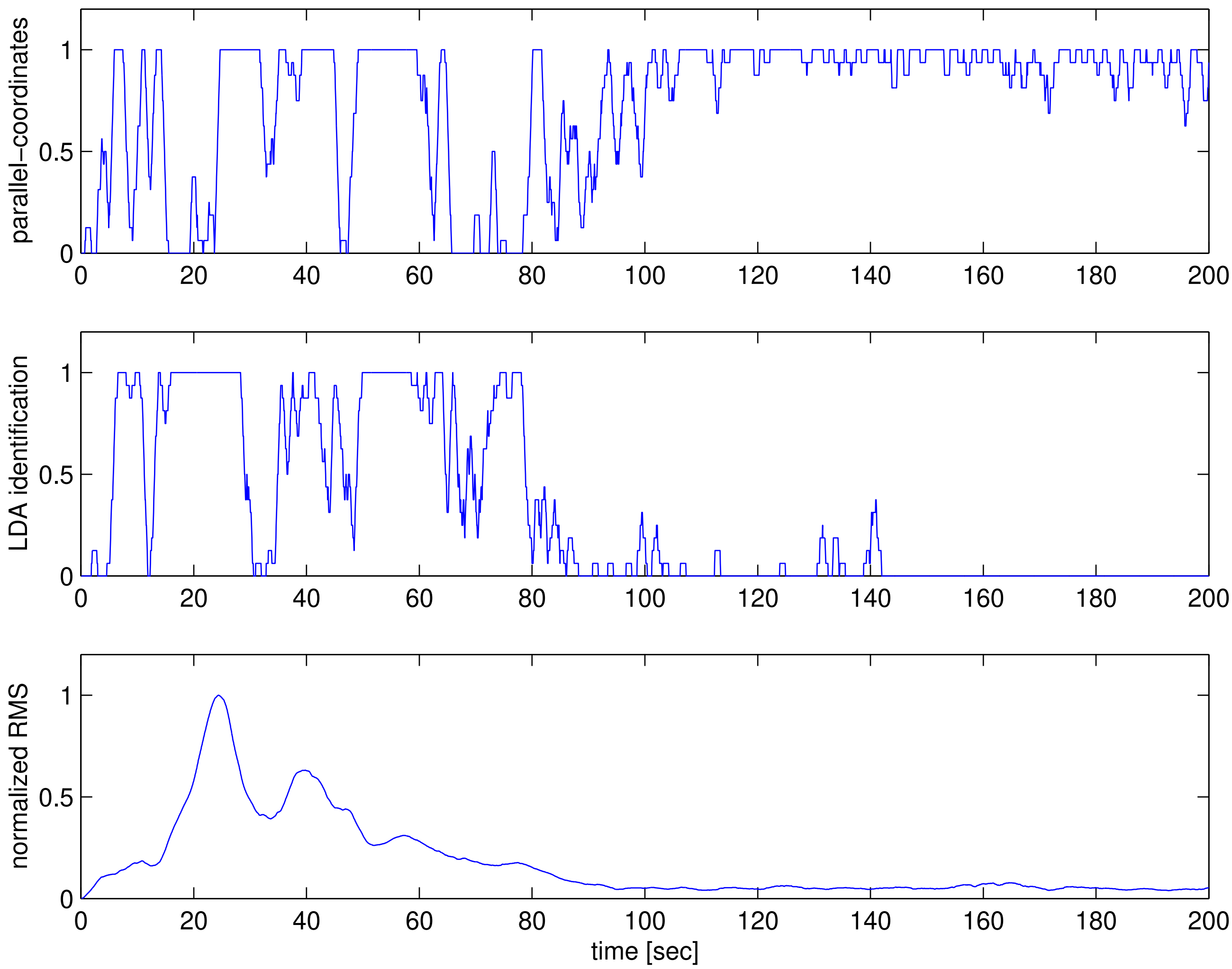

Figure 11: Identification of a G^1 vehicle. The bottom plot presents the normalized RMS of the acoustic signal. The top plots present the LDA and the parallel-coordinates identification performance. All three plots are synchronized in time. It appears that the parallel-coordinates classifier is capable of identifying the vehicle at larger distances than the LDA method.

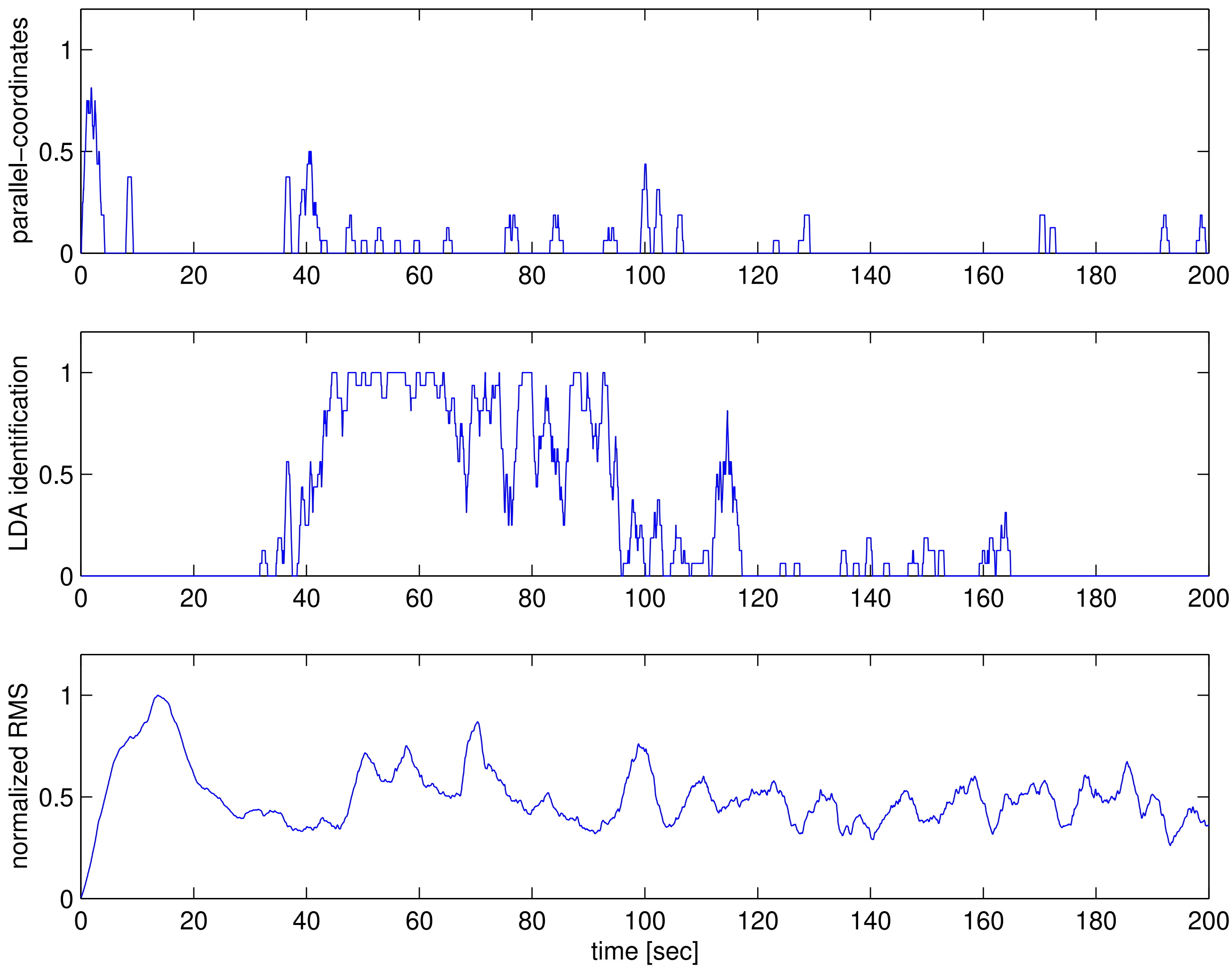

Figure 12: Identification of a G^2 vehicle. It appears that while the LDA suffers false identification in this recording, the parallel-coordinates classifier succeeds in discriminating this vehicle.

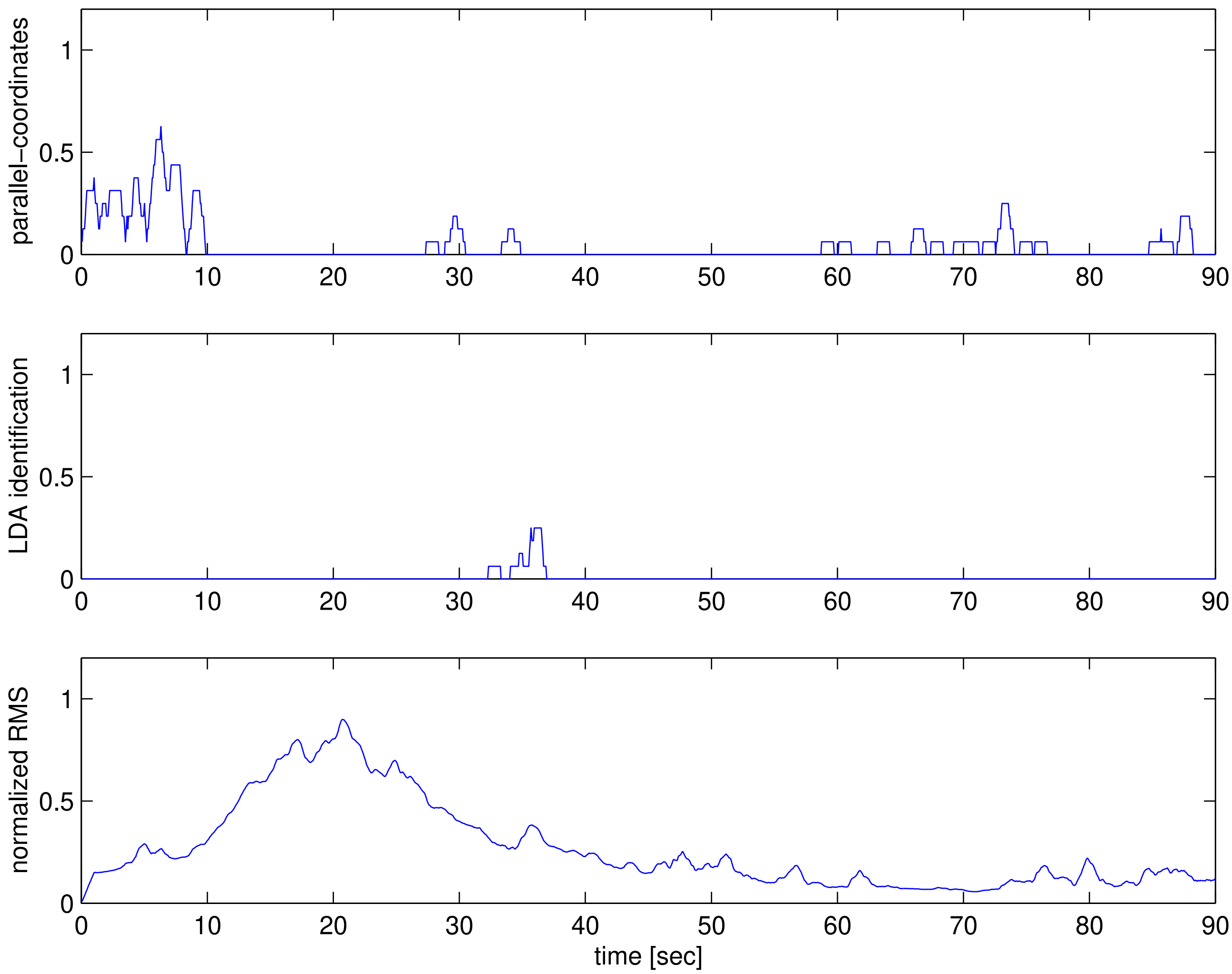

Figure 13: Identification of a G^3 vehicle. It appears that both methods discriminate this vehicle well.

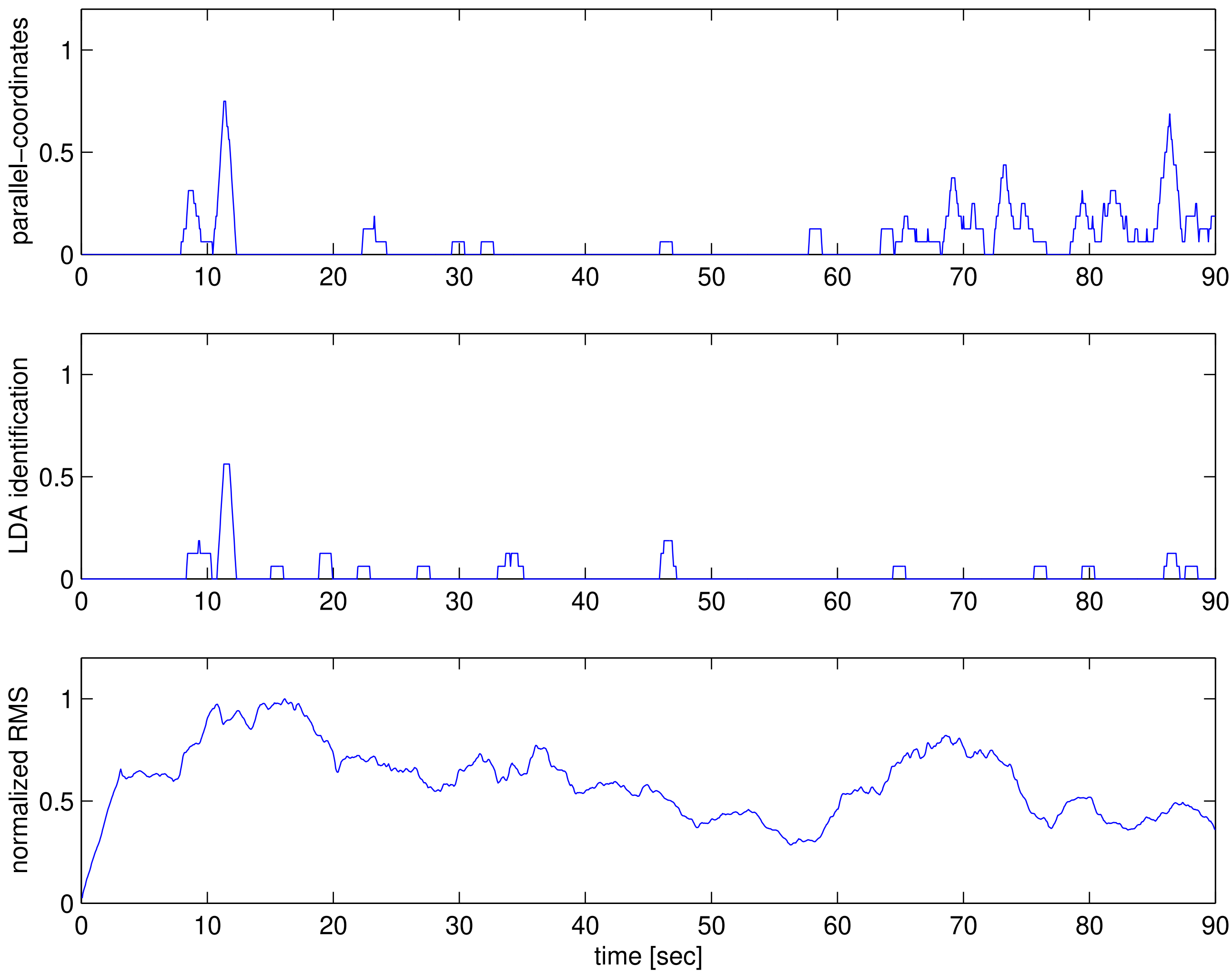

Figure 14: Identification of a G^4 vehicle. It appears that both methods discriminate this vehicle well.

Summary of the LDA identification results: It can be seen that the identification of the G^1 vehicle is successful during the first 90 seconds, and then fails. It can also be seen that the vehicles G^3 and G^4 are discriminated easily, but unfortunately the vehicle G^2 is falsely identified as G^1 in a large portion of the recording.

Summary of the parallel-coordinates identification results: It can be seen that the identification of the G^1 vehicle is as successful far away from the receiver as it is in its vicinity (the first 90 seconds). We stress, that, unlike LDA, the vehicle G^2 is properly classified as non-C^1. It is clear that the acoustic signature derived by the parallel coordinates both includes more acoustic modes of vehicle G^1 and less modes of different vehicles. This is due to the flexibility while defining the acoustic signature.

4.2 Open set identification

In order to test the acoustic signatures, various recordings of different vehicles were sliced and identified. The signature which was defined using the training set of the 4 vehicles $G^l,\ l = 1, 2, 3, 4$ remained unchanged, and a series of unfamiliar vehicles were presented. By unfamiliar we mean vehicles of types different than $G^l,\ l = 1, 2, 3, 4$. We label unfamiliar vehicles by $G^i \quad i \geq 5$.

Figures 15 and 16 present identification results from recordings of convoys of vehicles. These convoys contain one vehicle of type G^1 and a series of unfamiliar vehicles. Figures 17 and 18 present identification results from recordings of unfamiliar vehicles' convoys. The results show that the vehicles from class G^1 are identified, and none of the unfamiliar vehicles is falsely identified.

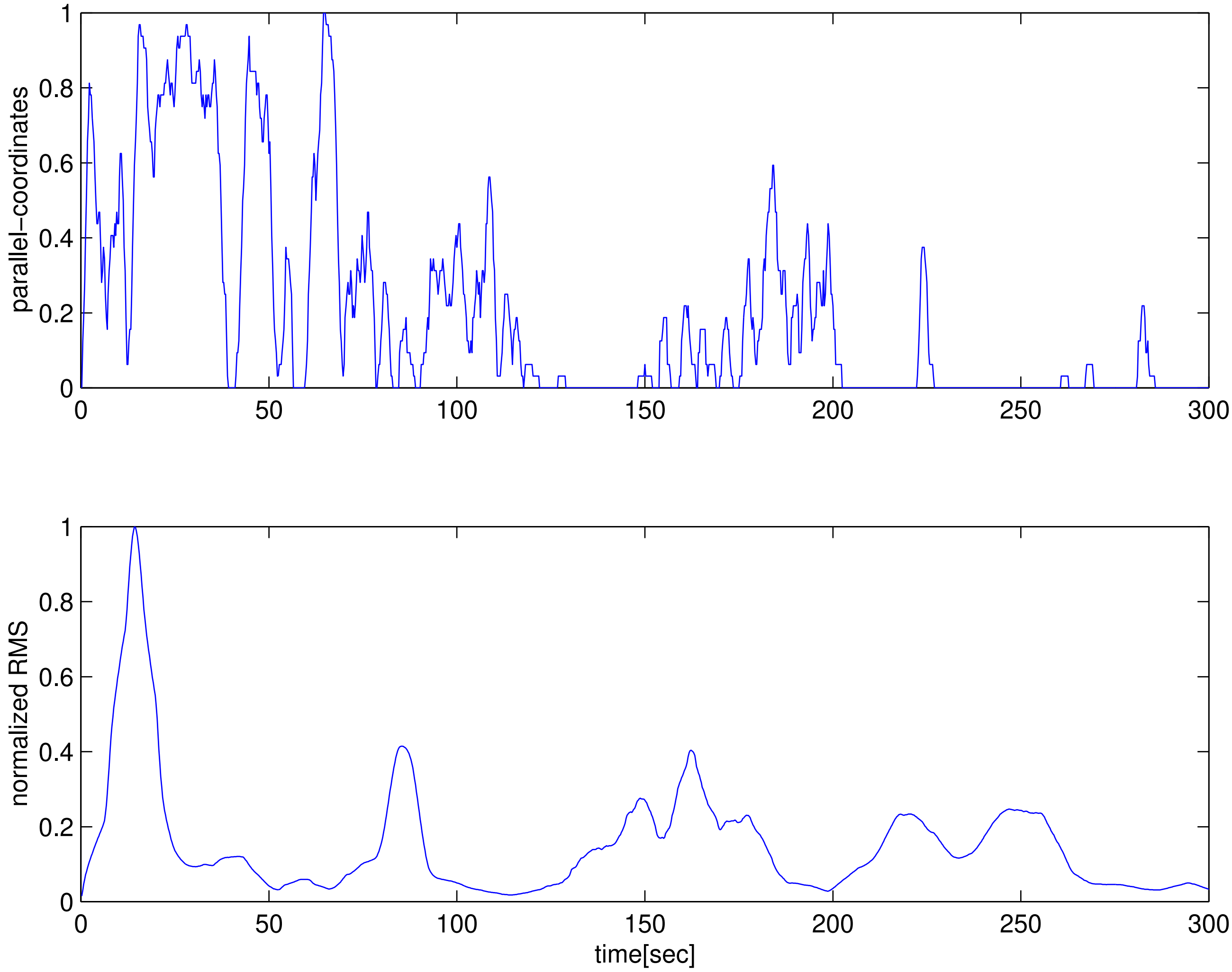

Figure 15: Identification of a convoy of vehicles. The first vehicle passing by the recording unit (at ≈ 20 seconds) is G^1. The following vehicles were not present in the training set (G^i for $i \geq 5$). Note that the G^1 vehicle is identified (at the begining of the recording), and none of the others is falsely identified.

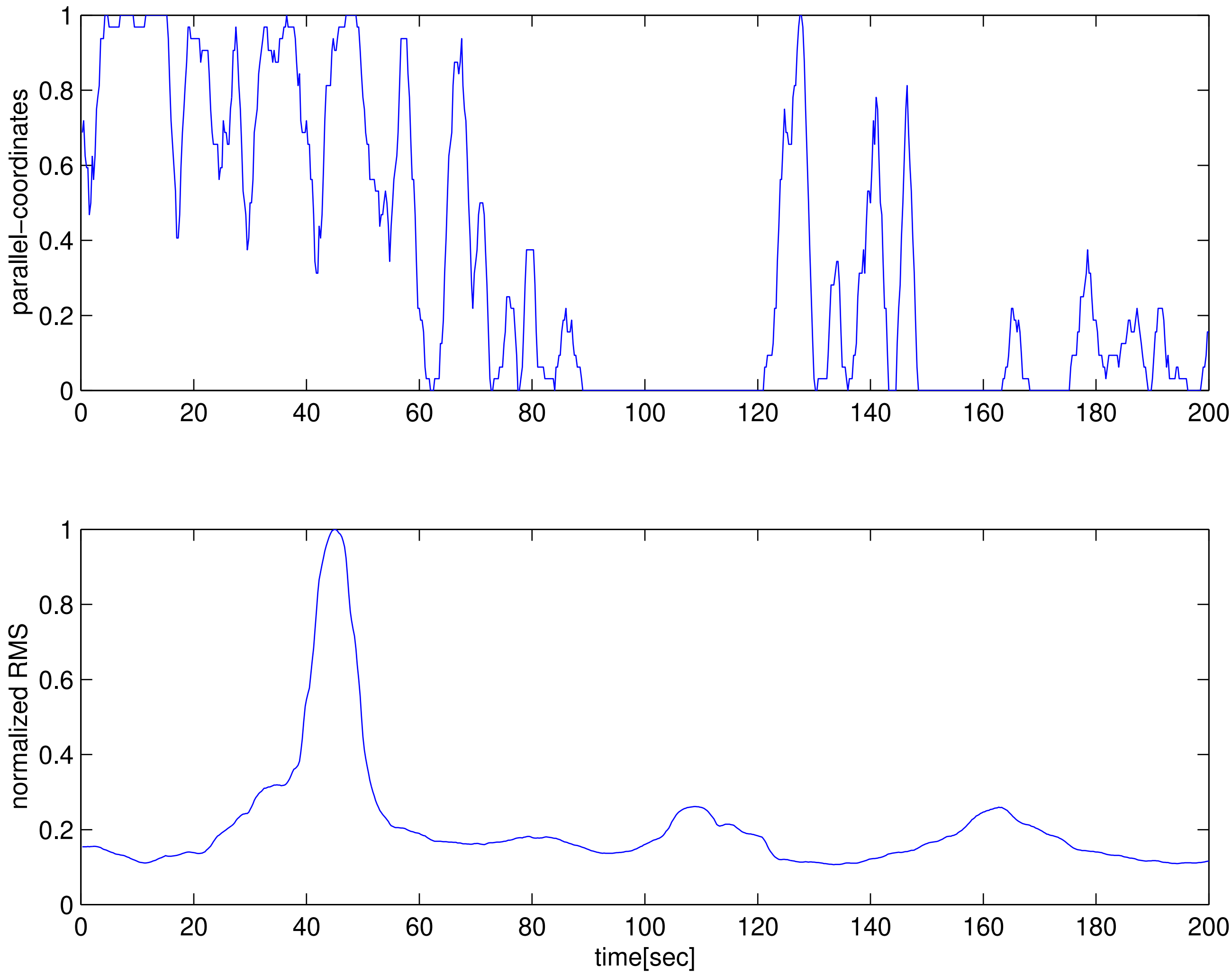

Figure 16: Identification of a convoy of vehicles. The first vehicle is from class G^1 (passing by the recording unit after 45 seconds). The rest of the vehicles were not in the training set (G^i for $i \geq 5$). Note that the G^1 vehicle is identified, and none of the others is falsely identified.

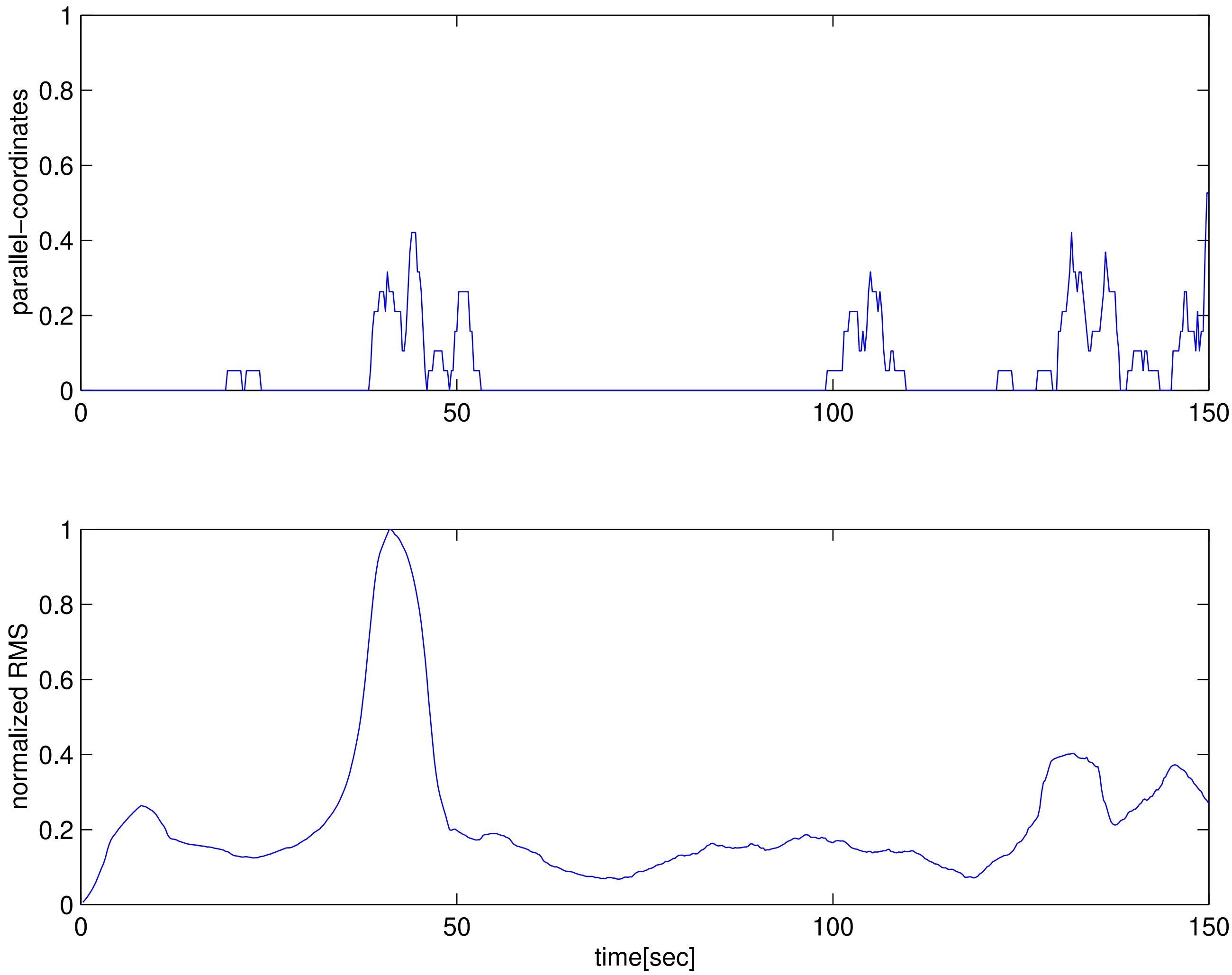

Figure 17: Identification of a convoy of vehicles. All the vehicles in the convoy are of types G^i, $i \geq 5$. Note that none of the vehicles is falsely identified.

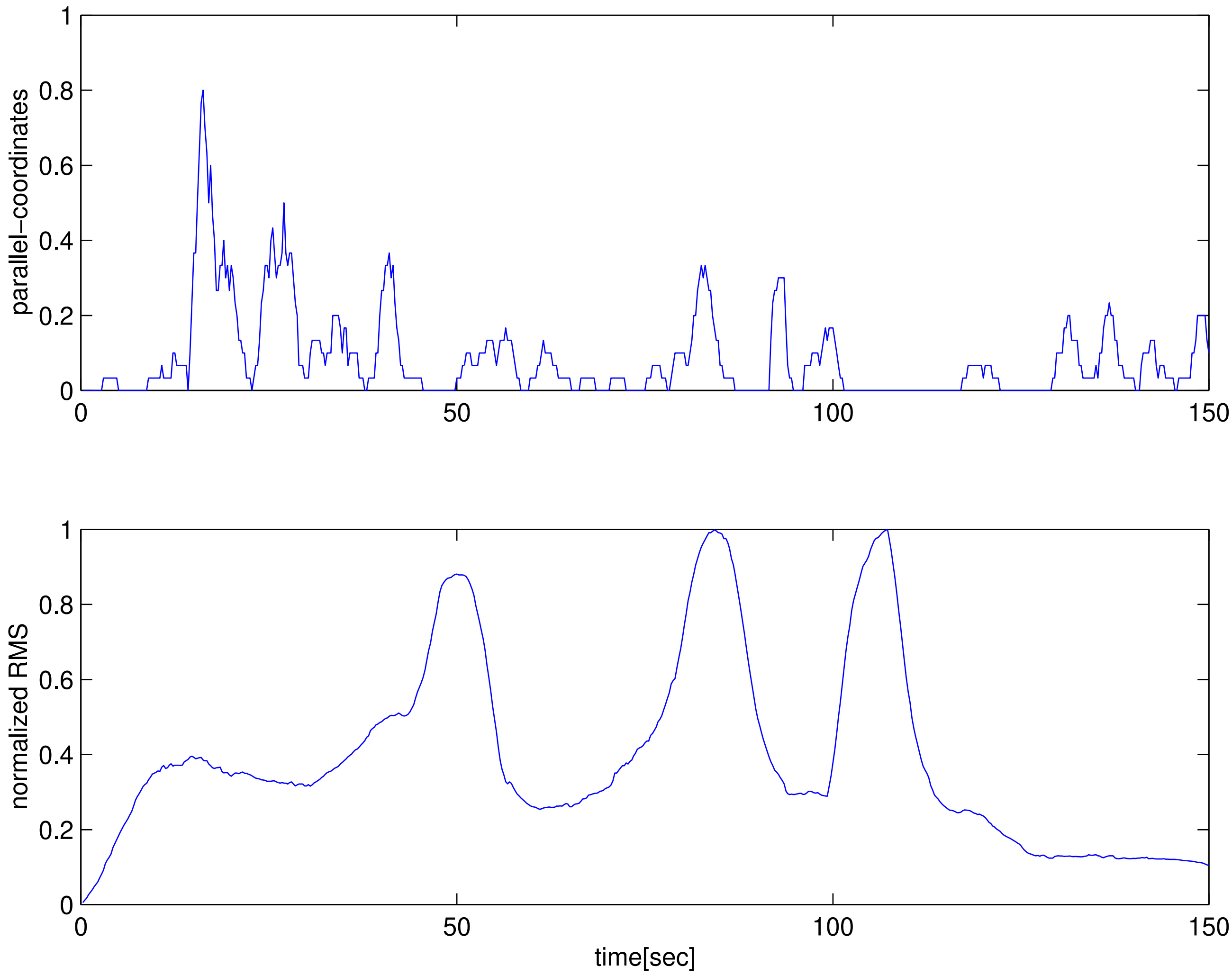

Figure 18: Identification of a convoy of vehicles. All the vehicles are of types G^i, $i \geq 5$. Note that none of the vehicles is falsely identified.

5 Conclusions

We described a method to extract an acoustic signature of a vehicle. The signature consists of several acoustic features unique to the vehicle. In our method, we choose cosine packets coefficients in the frequency domain, and obtain a multiscale decomposition of the acoustic signal. Then, we use the local discriminant basis to identify the features (frequency bands) that enable to separate the sought after vehicle from a group of other vehicles.

A vector of feature coefficients is extracted for each time slice of a recording. A training set containing a number of such vectors for each vehicle is gathered. The signature is defined on this training set using the parallel coordinates methods by a union of boxes and excluded holes. This union contains as many vectors originated from the sought vehicle and as few vectors as possible from the other vehicles.

The strength of the signature (e.g. the identification ability of the vehicle of interest, and the separation ability from different vehicles) is tested in two stages: 1. in a closed set, where the training set is constructed using 4 types of vehicles, and the identification test is performed on the same vehicles. 2. in an open set, where the signature defined in the first stage is tested against a series of unfamiliar vehicles.

The results show that we succeeded in extracting discriminating features and in defining an acoustic signature using parallel coordinates analysis. In turn, this signature performed well in identifying the vehicle of interest and in rejecting even unfamiliar vehicles.

This technology, which has many algorithmic variations, can be used to solve wide range of classification and detection problems which are based on acoustics processing and, more generally, for classification and detection of various signals which have near-periodic structure.

References

[1] A. Z. Averbuch, E. Hulata, V. A. Zheludev, I. Kozlov *A wavelet packet algorithm for classification and detection of moving vehicles*, Multidimensional Systems and Signal Processing, **12**, (2001) 9-31.

[2] A. Averbuch, V. Zheludev *Construction of biorthogonal discrete wavelet transforms using interpolatory splines*, Tel Aviv University, The School of Mathematical Sciences, Technical Report, (1999).

[3] A. Averbuch, F. G. Meyer, J.-O. Strömberg, Fast adaptive wavelet packet image compression, *IEEE Trans. Image Proc.*, **9**(5), 792-800, 2000.

[4] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees, Chapman & Hall, Inc., New York, 1993.

[5] P. Auscher and G. Weiss and M.V. Wickerhauser, Local Sine and Cosine Bases of Coifman and Meyer and the Construction of Smooth Wavelets, in *Wavelets-A Tutorial in Theory and Applications*, pp. 237-256, 1992.

[6] D. Donoho and I. Jonstone, Ideal denoising in an orthonormal basis chosen from a library of bases. *C.R. Acad. Sci. Paris,* Série I, 319, 1317-1322, 1994.

[7] J. Buckheit, D. Donoho, Improved linear discrimination using time-frequency dictionaries,*Proc. SPIE,* **2569**, (1995), 540-551.

[8] R. A. Fisher, The use of multiple measurements in taxonomic problems, *Ann. Eugenics***7**, (1936), 179-188.

[9] R. R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier a fenêtre,* C. R. Acad. Sci., pp. 259-261, 1991.

[10] G. Matviyenko, *Optimized local trigonometric bases*, Applied and Computational Harmonic Analysis, vol. 3, pp. 301-323, 1996.

[11] A. Inselberg, *Visual data minint with parallel coordinates* Computational Statistics, vol. 13, no. 1, pp. 47-63, 1998.

[12] A. Inselberg B. and Dimsdale, *Multidimensional lines. 1. reperesentation, 2. proximity and applications* SIAM Jour. on applied mathematics, vol. 54, no. 2, pp. 559-577, 578-596, 1994.

[13] A. Inselberg, *PARALLAX software package, www.math.tau.ac.ilãiisreal*

[14] R. R. Coifman, Y. Meyer, M. V. Wickerhauser, Adapted waveform analysis, wavelet-packets, and applications, *In Proceedings of ICIAM'91,* SIAM Press, Philadelphia, 1992, 41-50.

[15] R. R. Coifman, private communication, 1998.

[16] Q. Jiang, S.S. Goh, Z. Lin, Local discriminant time-frequency atoms for signal classification, *Signal Processing*, **72**, (1999), 47-52.

[17] R. R. Coifman, and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inf . Theory*, **38** (1992), 713-719.

[18] W. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software,* AK Peters, Wellesley, Massachusetts, 1994.

[19] S. Mallat, A wavelet tour on signal processing, Acad Press, 1998.

[20] S. Mallat and Z. Zhang, Matching pursuit with time-frequency dictionaries, *IEEE Trans. Sign. Proc.*, **41**(12) (1993), 3397-3415.

[21] N. Saito, *Local feature and its application using a library of bases*, PhD thesis, Yale University, New Haven, 1994.

[22] N. Saito and R.R. Coifman, *Local discriminant bases and their application*, J. Mathematical Imaging and Vision, vol. 5, pp. 337-358, 1995.

[23] N. Saito and R.R. Coifman, *Improved local discriminant bases using probability density estimation*, Proc. Am. Statist. Assoc., Statistical Computing Section, vol. 5, pp. 312-321, 1996.

[24] R. E. Karlsen, G. R. Gerhart, D. Gorsich and H. C. Choe, Wavelet-based ground vehicle acoustic recognition system, *Proceedings Seventh Annual: "Ground Target Modeling and Validation Conference', August 1996,* 249-256.

[25] I. Daubechies, Ten lectures on wavelets. SIAM, 1992.