International Journal of Wavelets, Multiresolution
and Information Processing
Vol. 2, No. 1 (2004) 1–22

IDENTIFICATION OF ACOUSTIC SIGNATURES FOR VEHICLES VIA REDUCTION OF DIMENSIONALITY

AMIR AVERBUCH*, EYAL HULATA and VALERY ZHELUDEV

School of Computer Science,
Tel Aviv University, Tel Aviv 69978, Israel
**amir@cs.tau.ac.il*

INNA KOZLOV

Electro-Optics Research and Development Ltd (EORD),
Technion, Haifa 32000, Israel

Received 16 September 2003

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 of 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 minimal 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 in the frequency domain of the acoustic signal. The proposed algorithm is robust even under severe noise and diverse rough surrounding conditions. This is a generic technology, which has many algorithmic variations, can be used to solve wide range of classification and detection problems which are based on a unique derivation of signatures.

Keywords:

1. Introduction

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

A crucial factor in having a successful classification (no false alarms) is to construct reliable signatures of classes that contain the right unique characteristic features that enable to discriminate between them.

Multiscale wavelet analysis provides a relevant methodology for feature extraction. However, most of the existing wavelet based techniques (for example Refs. 4, 18 and 19) lack translation invariance in the time domain. This is critical for detection of moving objects. Since misalignment between different signals can generate wrong results. Therefore, we developed in Refs. 2 and 14 a new approach which we call *Discriminant Block Pursuit*. 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 blocks of the wavelet packet coefficients of the signals. It is justified by the fact that each part of the vehicle emits a distinct acoustic signal, which in frequency domain contains only a few dominating bands. As the car moves, the conditions are changed and the configuration of these bands may vary, but the general disposition remains. Therefore, the blocks of the wavelet packet coefficients, each of which is related to a certain frequency band, are the relevant tool to base the classification on. As a decision unit in Ref. 2 we used a modified version of the Classification and Regression Tree (CART) algorithm.[3]

Another way to automatically reveal these characteristic frequency bands, which we pursue in the paper, consists of application of the multiscale local cosine transform (LCT)[5,6,15,21] in the frequency domain of the recorded signal. The blocks of the LCT coefficients, each of which is related to a certain frequency band, contain 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.[18–20] Note that the LCT method provides more flexibility in comparison to the method that is based on wavelet packets.

In order to reach a decision on the membership of an examined signal to a predetermined class, we use in the final phase of the process two classifiers: 1. A conventional classifier that is based on Linear Discriminant Analysis (LDA).[8] 2. A classifier, which is based on the "parallel coordinates" concept devised by Inselberg.[9–11] It is a visualization tool that automatically enables one to reduce the dimensionality of the problem. In all our experiments the parallel coordinates classifier overwhelmingly outperformed the LDA classifier. It proved to be a powerful tool for the creation of a meaningful signature for a class of signals.

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 close to zero.

The paper is organized as follows. In Sec. 2 we describe the structure of the acoustic signals we deal with and briefly outline the local cosine transforms. Then, we formulate our algorithm. In Sec. 3 we describe the algorithm which is centered around two basic issues: Selection of the discriminant blocks of the LCT and discrimination of signals. Then, we explain its implementation. In Sec. 3.3.2 we explain building the classifier based on the "parallel coordinates" methodology. Section 4 presents the experiments.

2. Formulation of the Approach

2.1. *The structure of the signals*

The raw data that were used to train the algorithm is a collection of acoustic recordings of four types of vehicles. The goal of this work is to obtain a unique identification of the vehicle labelled G^1. As a reference set we used recordings that were taken from three different vehicles labelled $G^{2,3,4}$. The signal from each vehicle was recorded while travelling at typical driving velocities. The data were collected in field conditions, while the different vehicles travelled toward and away from the recording unit up to a distance of 1,500 m. The acoustic signal was processed by a linear phase analog filter (cutoff at 512 Hz) and then digitized at about 1 kHz (1,024 samples per second) by a 12-bit digital recorder. This cutoff frequency was selected since the unique acoustic tones emitted by the vehicle are all within the range of 0 to 512 Hz.

The basic time fragment for the analysis is called *slice* and labelled by S. Each slice is $2^8 = 256$ samples long (0.25 seconds). The time interval is selected to be 0.25 of a second since the acoustic signal is considered to be quasiperiodic. 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 4 Hz width. Figure 1 presents time slices of acoustic recordings for each of the four vehicles. Figure 2 presents the spectra of these slices.

We can see from Fig. 2 that each of the vehicles possesses a unique set of acoustic frequencies (or frequency events). We treat these frequency events as acoustic features.

In order to better visualize the features, Fig. 3 presents a spectral map of the recordings of the vehicle in G^1. The map is constructed by the application of the Fourier transform on consecutive slices. Each slice is 0.25 time duration, 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). We can observe that the location of each feature varies with time within a certain frequency band. Following the suggestion in Ref. 7 we increase the number of slices in the training sets and in the test sets by allowing adjacent slices to overlap with each other. By this means we also enhance the robustness of the identification.

The problem, which we solve, can be formulated as a two-class classification problem where

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

Class C^2 contains signals emitted by different type of vehicles.

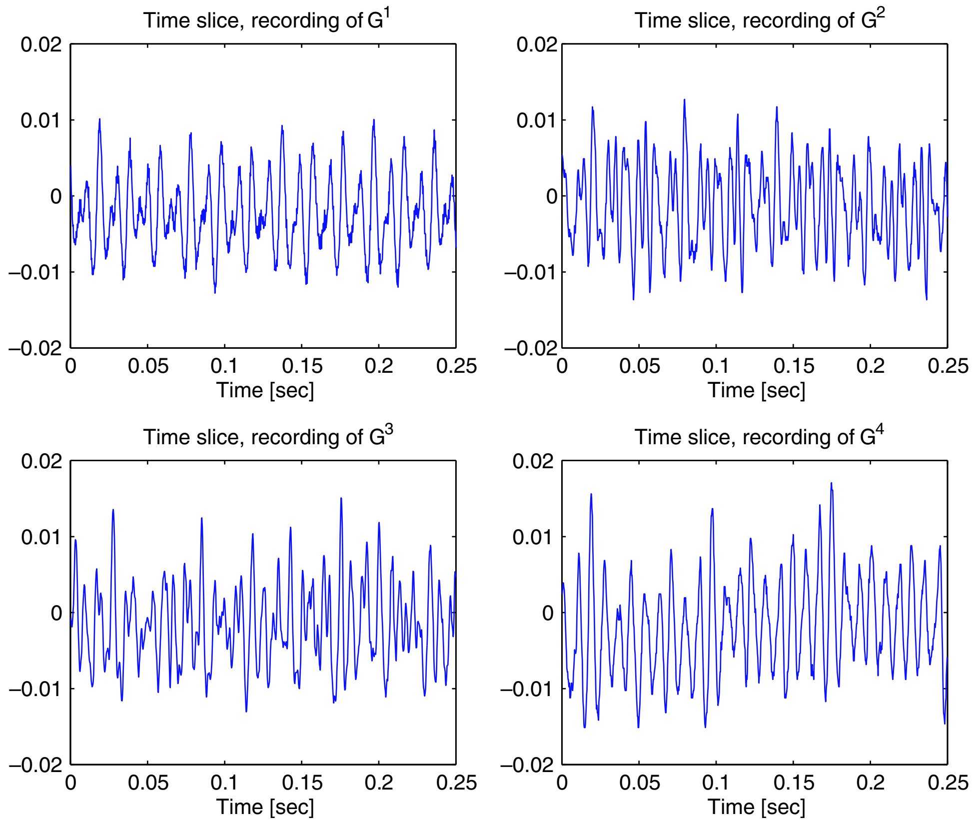

Fig. 1. Time slices from the four vehicles.

2.2. *The cosine packet transform*

To derive a signature of a class of signals we need an efficient tool that performs an automatic adaptive search for valuable frequency bands of a signal or a class of signals. One way to extract these characteristic bands is to apply the wavelet packet transform in the time domain of the signal and to use as characteristic features energies in the blocks of wavelet packet coefficients.[2]

In this paper we explore another way, which, to some extent is similar to the previous one. Namely, we apply the local cosine transform (LCT)[5,15] analysis on the frequency domain of the signals. This is equivalent to presentation of the signal through linear combinations of the inverse Fourier transforms of the local cosines. Once implemented, the multiscale local cosine transform yields a variety of different partitions of the frequency domain. As explained in Ref. 1, we can adaptively select the sizes and locations of the windows using the best basis algorithm.

The cosine packet transform creates a binary tree of the Fourier spectrum of the signal. In the first step, the Nyquist frequency band is divided into two half-bands,

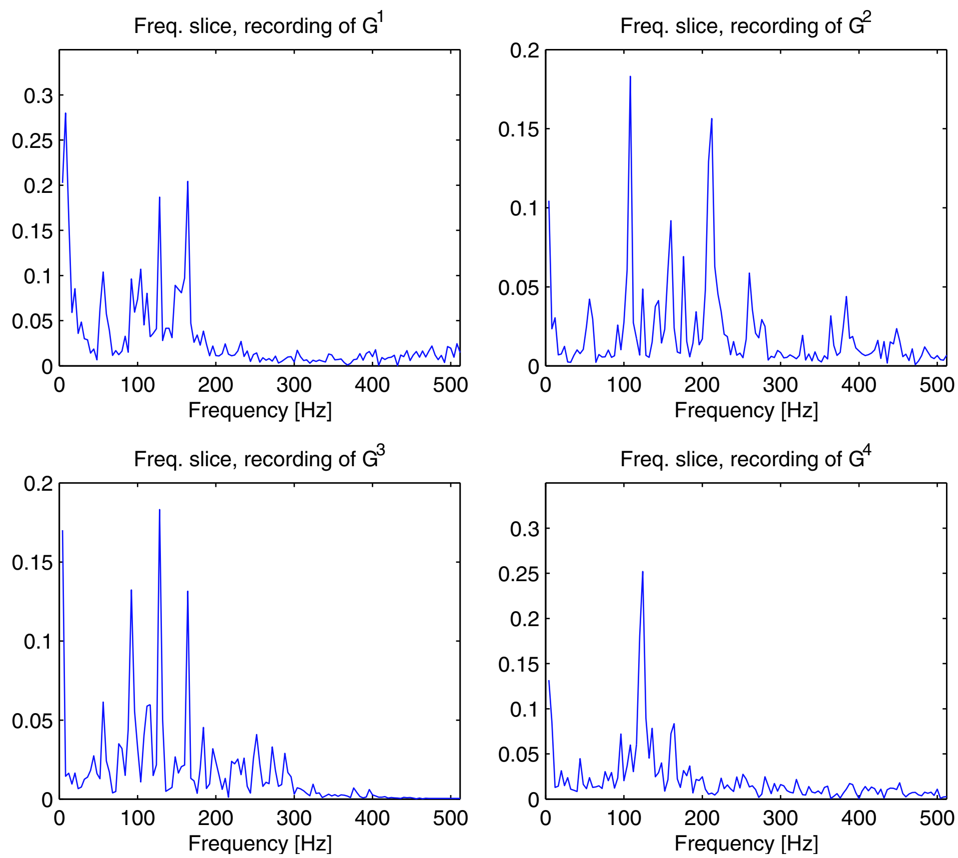

Fig. 2. Frequency domains of slices from the four vehicles.

and the discrete local cosine transform is applied to the spectra lying within each half-band. Then, we decompose further each half-band into two quarter-bands and apply the local cosine transform on each quarter-band. By recursive application of these procedures we obtain a homogeneous binary tree-structured decomposition as shown in Fig. 4.

Obviously, the two blocks of the transformcoefficients of the first level (half-bands) can be regarded as parents of the corresponding pairs of blocks of the second level (quarter-bands) and so on. Local cosines associated with a certain level form an orthogonal basis of the space of signals. But, introducing a cost function for the sets of transformcoefficients, we can design an orthogonal basis, which is optimal for a given signal or a class of signals with respect to this function. It is done by comparing the cost of parent blocks of coefficients with the cost of their children blocks and eliminating blocks with the higher cost. In our setting design of an optimal basis results in an optimal segmentation of the frequency band of a given

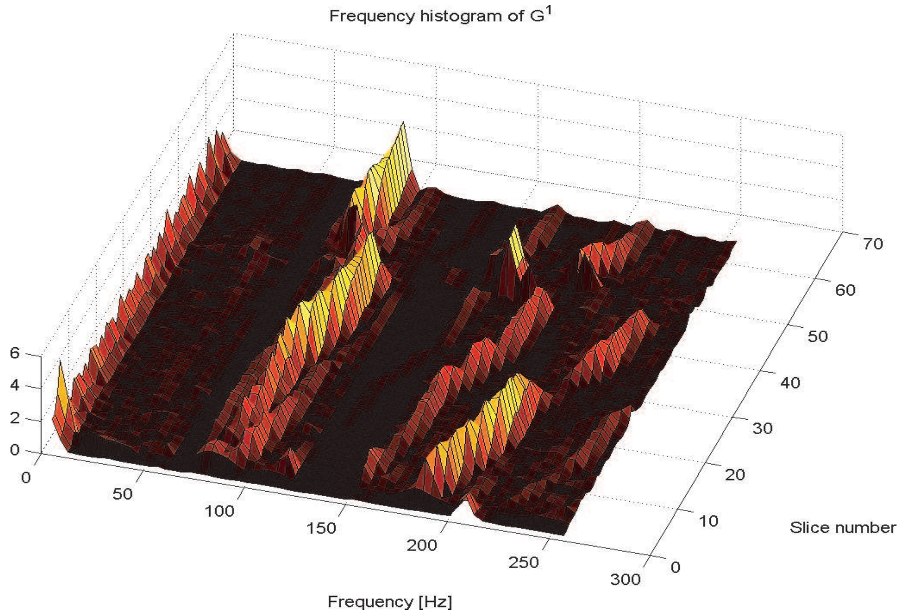

Fig. 3. Frequency map of the vehicle G^1. Each row in the figure is a frequency domain of one slice recording. Each slice is 0.25 second long in time and the presented frequency domain is 256 Hz wide (64 bins, 4 Hz per bin). The sliding ridges are the acoustic features (tonal lines) whose instantaneous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

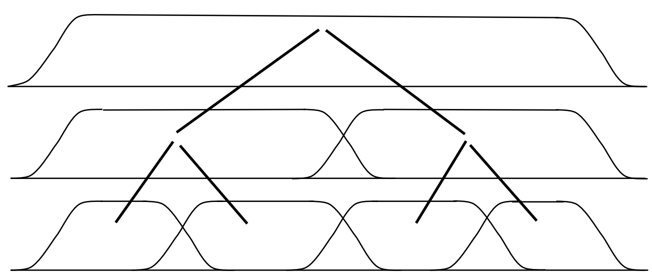

Fig. 4. Binary tree decomposition. Within each interval the signal is expanded into a local DCT. The optimal segmentation is then searched for.

signal or a class of signals. For our goal we should select a cost function, which enables one to design a basis, which optimally discriminates a sought for G^1-vehicle from vehicles belonging to other classes. We specify our cost function in Sec. 3. One of the important features of the transform is its computational efficiency. The local cosine transform is implemented via the fast DCT-IV transform using the operation of *folding* (see Ref. 15). Selection of the optimal basis is carried out automatically. Additional adaptation means can be added by the choice of windows for the LCT[17] and by manual segmentation of the frequency domain in order to better capture the characteristic frequency bands of a signal.

2.3. *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 local cosine transform (LCT) to the signal to be classified. We use as its characteristic features the l_2 norms of the LCT coefficients contained in the previously selected blocks.
(3) Finally, we submit the vectors of the extracted features to the "Parallel Coordinates" classifier, which was trained beforehand. It decides which class this signal belongs to.

3. Description of the Algorithm and its Implementation

3.1. *Skeleton of 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. Construction the classifier that is 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.[18–20]

Construction of the training set: Initially, we collect as many as possible recordings of the signals emitted by vehicle G^1 (Class C^1 signals) 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. Altogether 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:

1. The power spectrum of each slice F_i^l 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.
2. The multiscale local cosine transform is applied to the slices $\{F_i^l\}$ up to a level (scale) $m < 7$. This procedure produces mn coefficients for each slice $\{F_i^l\}$. These coefficients are 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. The subscript i is the index of the slice, the superscripts l and s designate the class the slice belongs to and the index of the block in the multiscale transform, respectively. The variable j indicates the position of a coefficient within the given block.
3. The energy map of the block $B^{l,s}$ is defined as the normalized mean square 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 number s, denoted by B^s, in the classes C^1, C^2. The discriminant power (DP) of this block is defined as the sum:

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

This DP can serve as a cost function for the selection of the most discriminating basis.

Selection of discriminating blocks: Now we are in the position to select a few discriminant blocks which form the acoustic signatures for the classes. Since we are in a situation where the supports of the parent and children blocks overlap producing an overcomplete coverage of the space, we use the *Best Basis Selection Algorithm*[6] to discover which block in the multiscale expansion has a stronger DP. The idea is to compare the DP of each pair of the "children" with the DP of their parent in the multiscale expansion. In the case when the DP of the parent exceeds the sum of the DP of the children, the children blocks are discarded and vice versa. Formally, this recursive construction can be defined by the following way. Each block of coefficients at a level n is indexed as $\mathbf{B}_n^p$, where p indicates its position from left to right. The set of blocks

$\{B_n^p\}_{p=0}^{2^n-1}$ of the level n corresponds to an orthogonal basis of the signal space. The set of blocks $\{{}_dB_n^p\}$ which corresponds to the "most discriminating basis" for this space is defined by the recursive rule (from bottom up along the tree branches).

$$ {}_dB_n^p = \begin{cases} B_{n+1}^{2p} \cup B_{n+1}^{2p+1} & \text{if } \mathbf{D}(B_{n+1}^{2p}) + \mathbf{D}(B_{n+1}^{2p+1}) > \mathbf{D}(B_n^p) \\ B_n^p & \text{otherwise.} \end{cases} $$

This non-overlapping set of blocks covers the whole frequency domain of our recorded signals. 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 ${}_dB^s, s = 1, \ldots, M$ 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 the highest discriminant power. In Fig. 5 we present the coefficients of

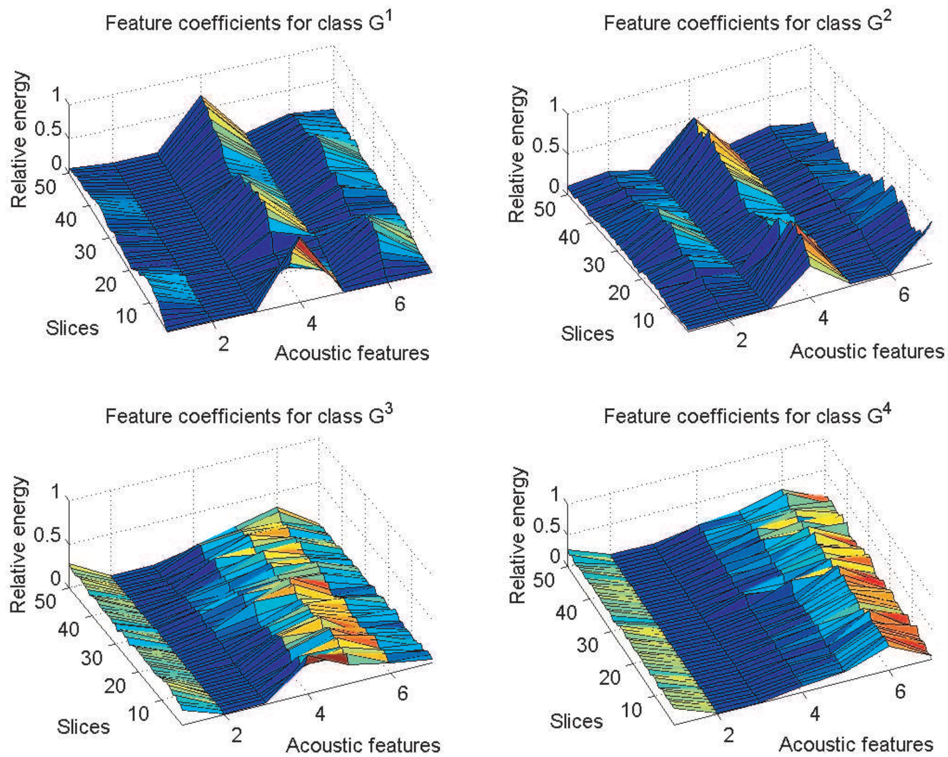

Fig. 5. Acoustic feature coefficients of the $G^{1,2,3,4}$ vehicles. For each vehicle consecutive slices were decomposed using LCT and the seven blocks were extracted and normalized. The coefficients of each slice are presented by a polygonal line, and a histogram of those values over time is built for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

those selected LCT blocks for four vehicles. The figure is presented similarly to the spectral map in Fig. 3. The difference is that instead of frequencies of subsequent slices, the blocks' coefficients are presented over time. Note that the sliding ridges as in Fig. 3 are transformed into relatively constant block values.

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 $B^s, s = 1, \ldots, M$, we proceed to the classification phase.

3.3.1. *Preparation of the reference set*

Initially, we choose a number of recordings with known membership in the classes $C^l, l = 1, 2$, from which we form the reference set. These recordings are sliced similarly to the recordings, which were used for the preparation of the training set. We form from all the selected recordings related to a certain class C^l a number of overlapping slices. Each is of length $n = 256$ and corresponds to 1 second time interval. These slices are shifted with respect to each other by 64 samples. We assume that there are μ^l slices $\{a^l(i)\}_{i=1}^{\mu^l}$, 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 a scale m to 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 ${}_dB^s, s = 1, \ldots, M$ that were selected before. The energy of a block ${}_dB^s = \{b^s(j)\}_{j=1}^{R_s}$ is defined as the average of squared magnitudes of its terms:

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

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

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

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

The reference matrices 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 and to exploit it to detect signals from this class with minimal false alarm. The representative vectors for slices originated from class C^1 are embedded into an M-dimensional hyperplane space. We have to find a domain in this hyperplane that contains as many vectors from class C^1 and as few vectors from the other class C^2 as possible. This domain is referred to as the *acoustic signature* of class C^1.

In this section, we present our method for identifying the acoustic signature. We do so by training a classifier using the reference set formed from the 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.[9–11] 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 Fig. 6 present in parallel coordinates seven-dimensional subspaces containing the signatures of vehicles of four types.

The problem then is reduced to framing a subset in the hyperplane, that contain vectors representing slices of the sought after vehicle and does not contain vectors related to other vehicles. Such a subset of the hyperplane 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 the coordinates of the box that easily separate the classes and so also the coordinates which have overlapping values for different classes. The acoustic signature box is 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 as possible is excluded. Figures 7–9 demonstrate the method (using Parallax[11]).

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 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

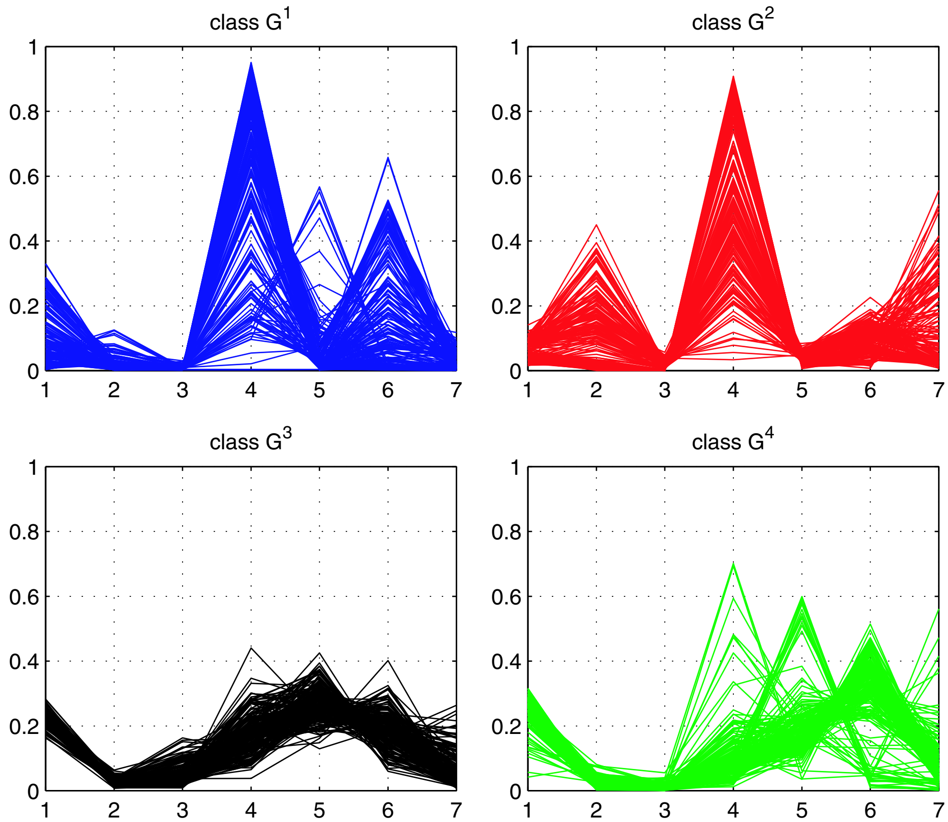

Fig. 6. Parallel coordinate plot in 7-D of the vectors $G^{1,2,3,4}$. Each class is presented in a different subplot. Each slice is transformed into a vector in 7-D (seven block coefficients). Such a vector is drawn as a polygonal line using seven parallel coordinates.

advantages of this method are:

1. The treatment of false identifications is simple. A non-C^1 vehicle is falsely identified if its representing set of vector overlaps with the union of boxes forming the signature of the Class C^1. 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 was not present in the training phase) is more likely to be classified correctly than while using LDA. Even in the case when an unknown

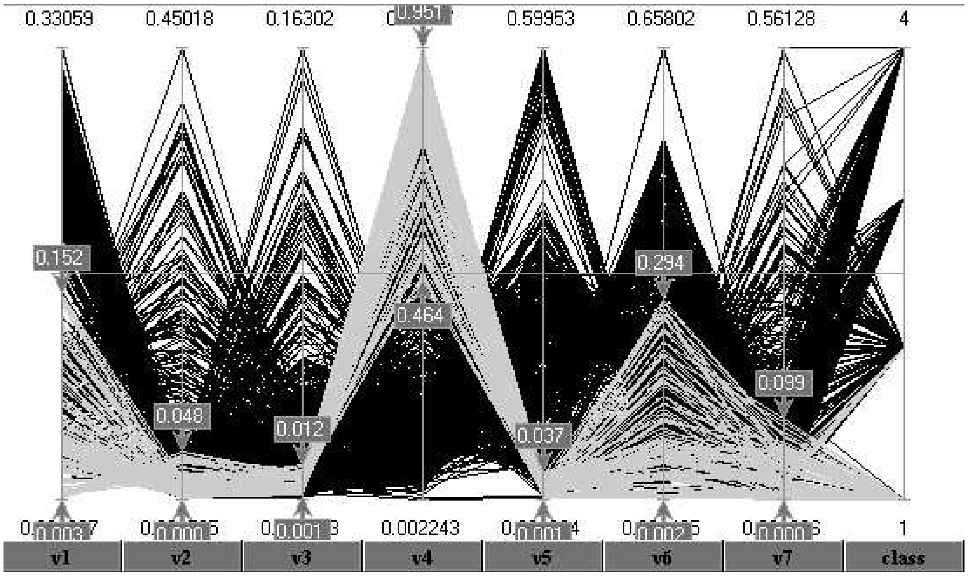

Fig. 7. Parallel coordinate plot in 7-D of the vectors of class C^1 (blue lines). The vectors of the other slices are in black. The seven coordinates are represented by the first seven columns. The final column (labeled *class*) separates the four types of vehicles($G^{1,2,3,4}$).

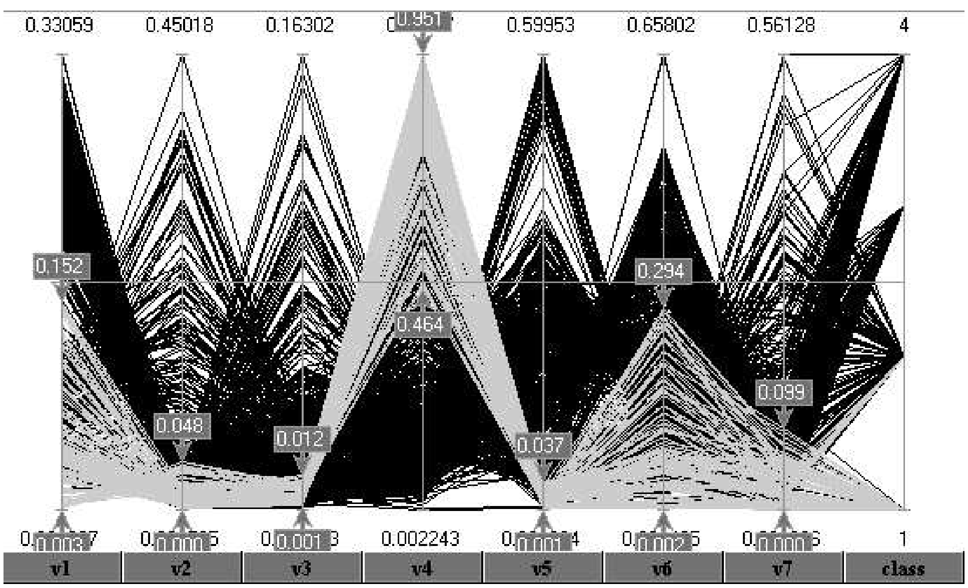

Fig. 8. A box framing class C^1. The blue lines represent vectors that are contained in the box, the other vectors are in black. The box (the bounds in each coordinates are marked by arrows) contains 60% of the vectors of class C^1, and a few vectors of type G^2. The little overlap with G^2 does not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

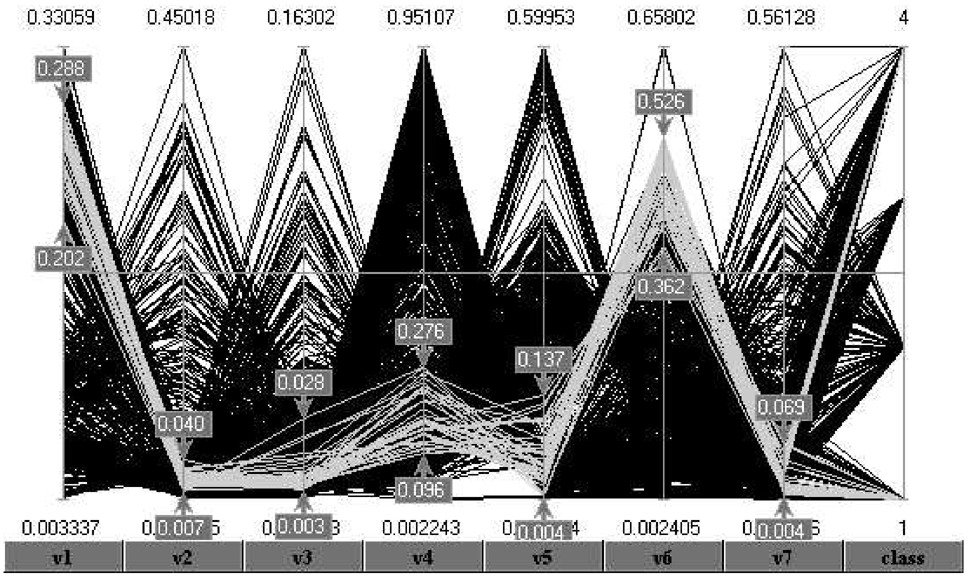

Fig. 9. Another box framing class C^1. The blue lines represent vectors that are contained in the box, the other vectors are black. The box (the bounds in each coordinates are marked by arrows) contains 20% of the vectors of type G^1 (that are not contained in the previous box), and very few vectors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

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 must be capable of processing 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 applied to each row $T(i,:)$ of this matrix up to the 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 detection 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 related to the signal S is ready, we present each row $W(i,:)$ of the matrix to two classifiers:

1. LDA calculates the Mahalanobis distances between the vector $W(i,:)$ and two 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 to check 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 output (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 is the domain below the curve, the better the identification is. Figure 10 demonstrates how this is done.

4. The Identification Results

We conducted two series of experiments:

1. **Closed set identification**, in which the four types of vehicles ($G^{1,2,3,4}$) participate. We compare performance of the LDA and the parallel-coordinates classifiers. These experiments demonstrate that the parallel-coordinate classifier is superior to the LDA in the separation of the classes.
2. **Open set identification**, in which a series of known and unknown vehicles are 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[a] sketch of the acoustic signal of the moving vehicle is presented. This sketch is synchronized with the identification results. Note that normally RMS attains its maximum when the

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

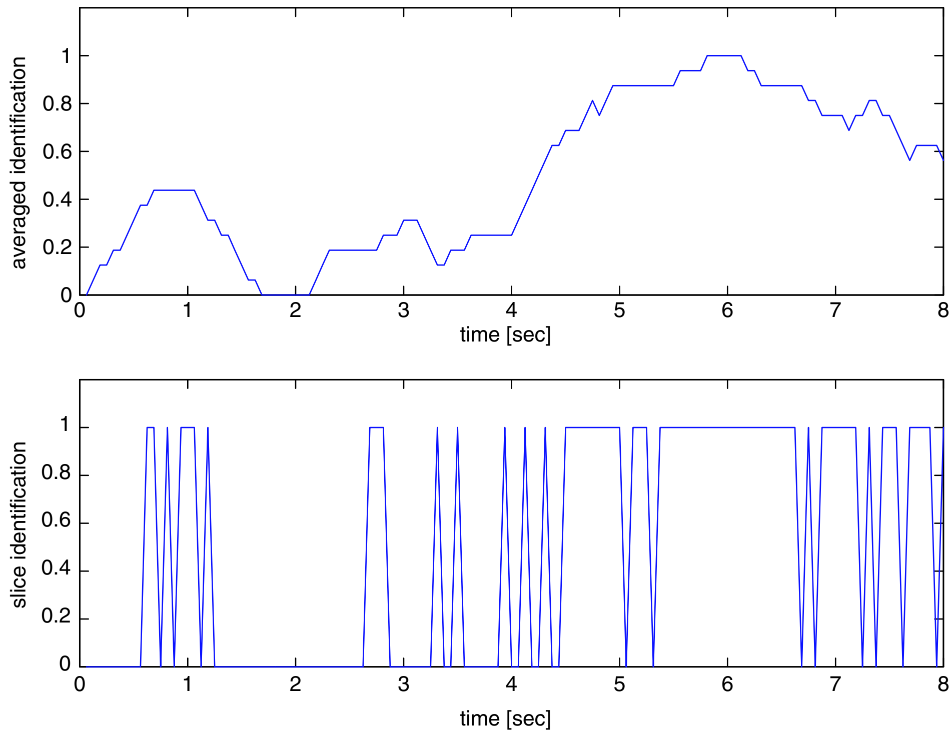

Fig. 10. An example of the identification process of a recording. The bottom figure shows the result of classification of the vectors $\{W(i,:)\}_{i=1}^{128}$. The top figure shows the output of a moving average of the bottom figure.

vehicle passes by the recording unit. As the the vehicle recedes from the recording unit, RMS fades.

4.1. *Closed set identification*

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

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.

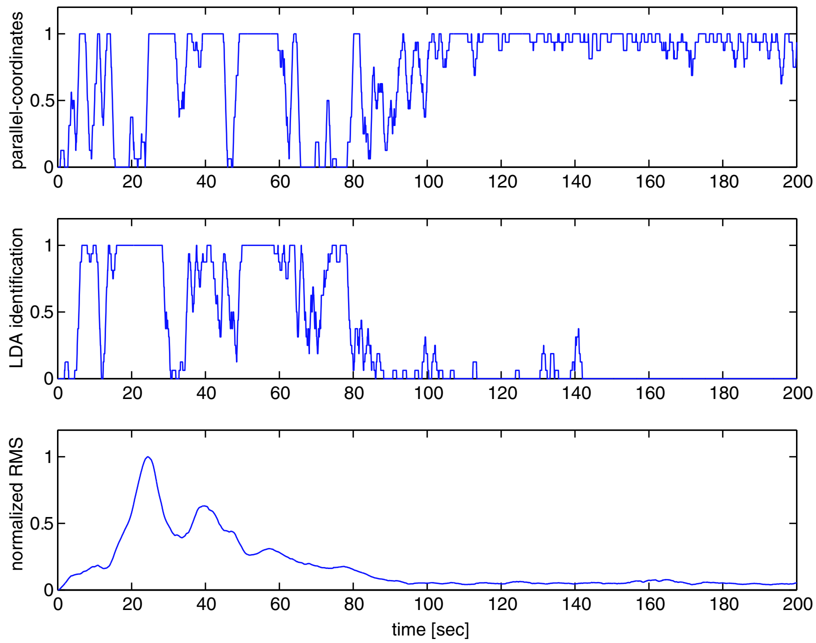

Fig. 11. Identification of a G^1 vehicle. The bottom plot presents the normalized RMS of the acoustic signal. The two upper plots present the LDA and the parallel-coordinates identification performance. All three plots are synchronized in time. It appears that the parallel-coordinates classifier is capable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

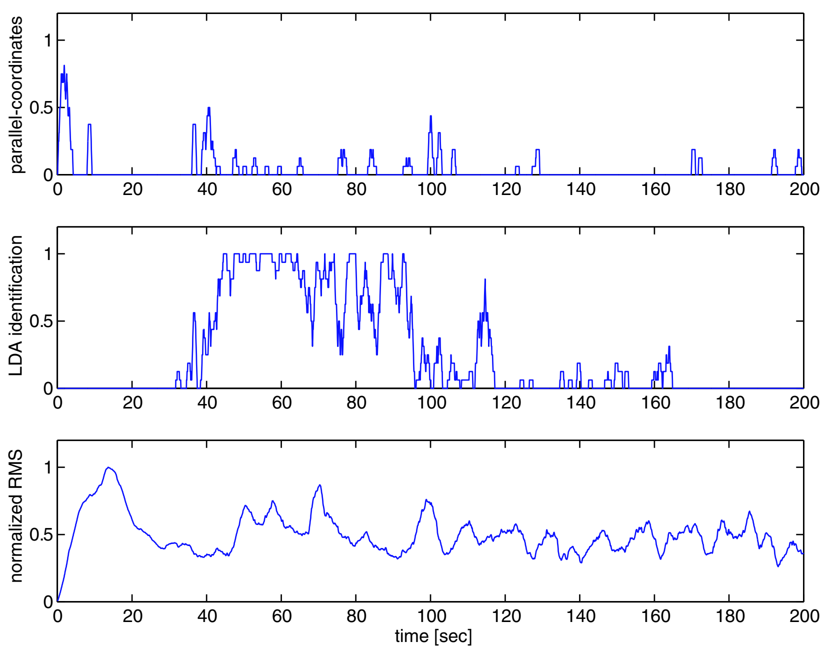

Fig. 12. Identification of a G^2 vehicle. It appears that while the LDA suffers false identification in this recording, the parallel-coordinates classifier succeeds in discriminating this vehicle.

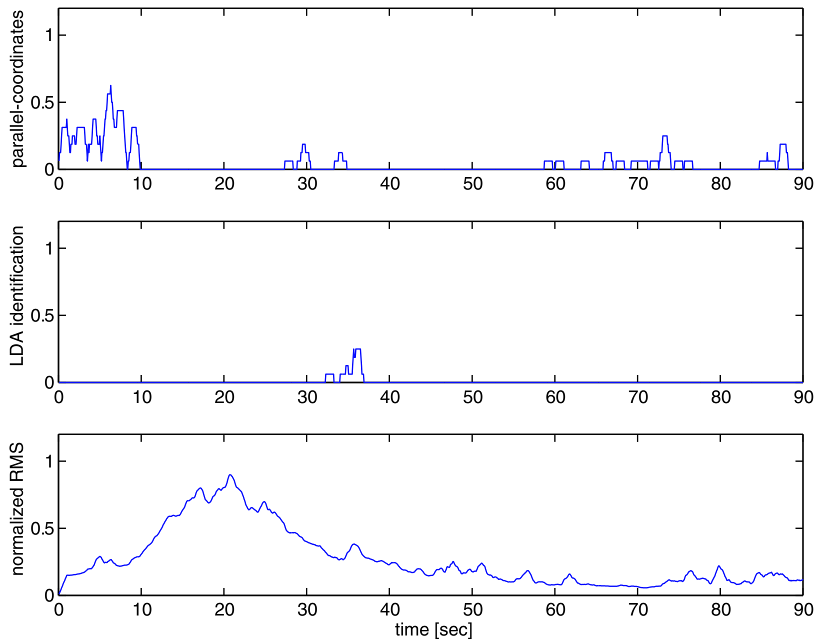

Fig. 13. Identification of a G^3 vehicle. It appears that both methods discriminate this vehicle well.

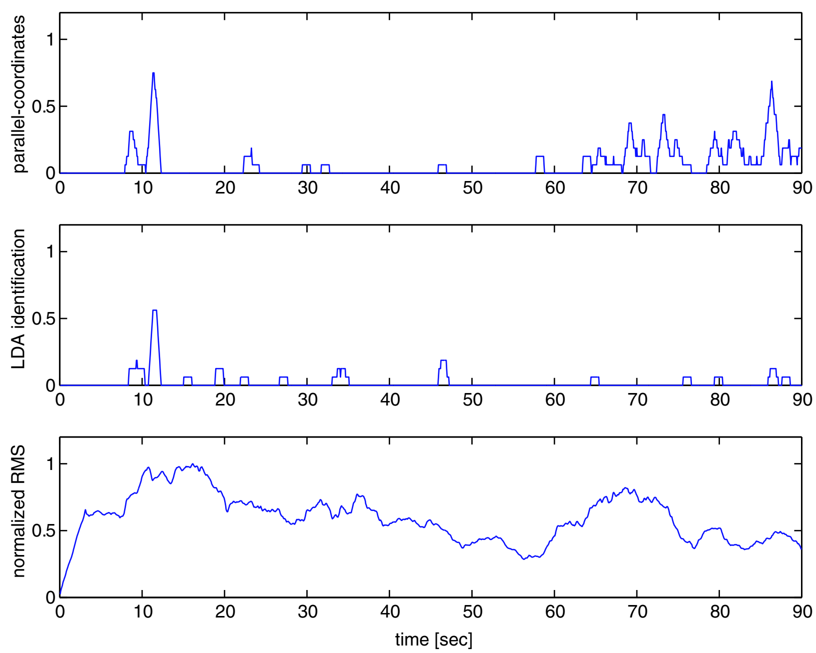

Fig. 14. Identification of a G^4 vehicle. It appears that both methods discriminate this vehicle well.

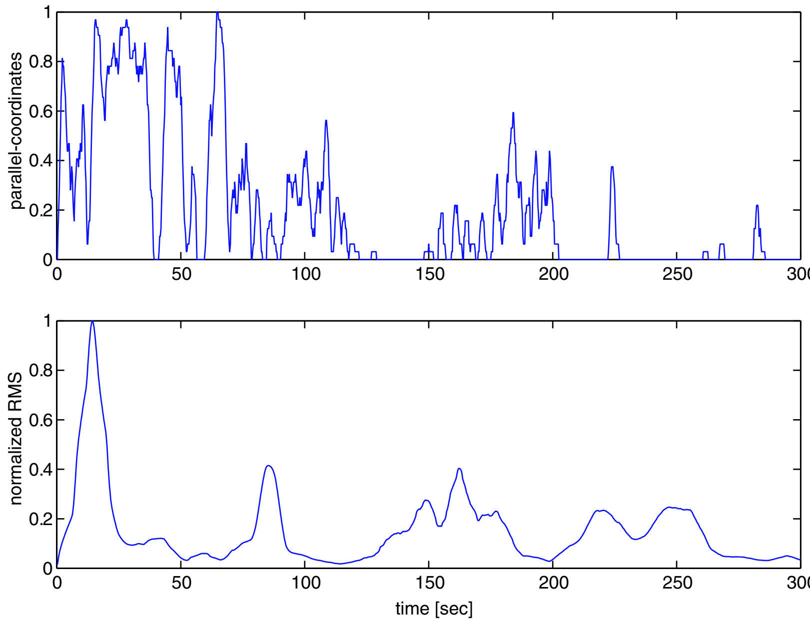

Fig. 15. Identification of a convoy of vehicles. The first vehicle passing by the recording unit (at ≈20 seconds) is G^1. The following vehicles were not present in the training set (G^i for $i \geq 5$). Note that the G^1 vehicle is identified (at the beginning of the recording), and none of the others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

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 four 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 from $G^l, l = 1, 2, 3, 4$. We label unfamiliar vehicles by $G^i, i \geq 5$.

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.

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

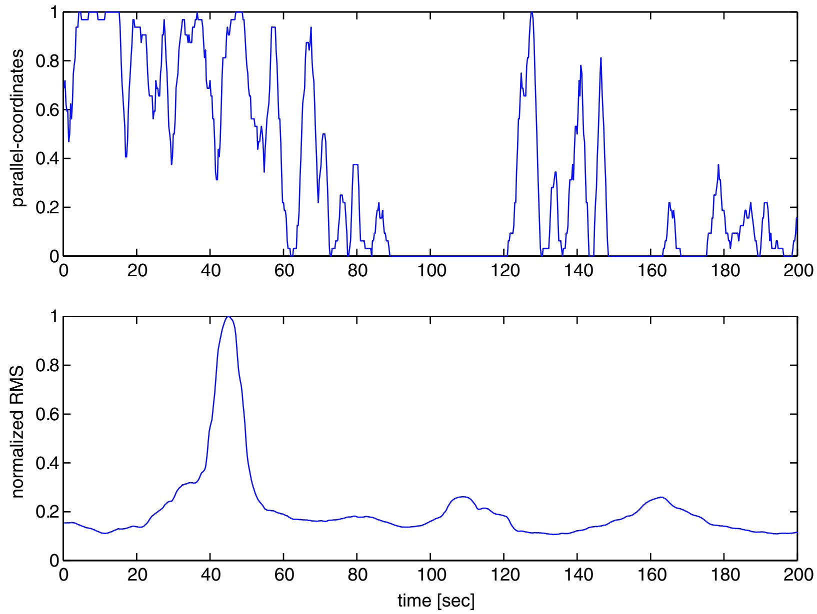

Fig. 16. Identification of a convoy of vehicles. The first vehicle is from class G^1 (passing by the recording unit after 45 seconds). The rest of the vehicles did not participate in the training set (G^i for $i \geq 5$). Note that the G^1 vehicle is identified, and none of the others is falsely identified.

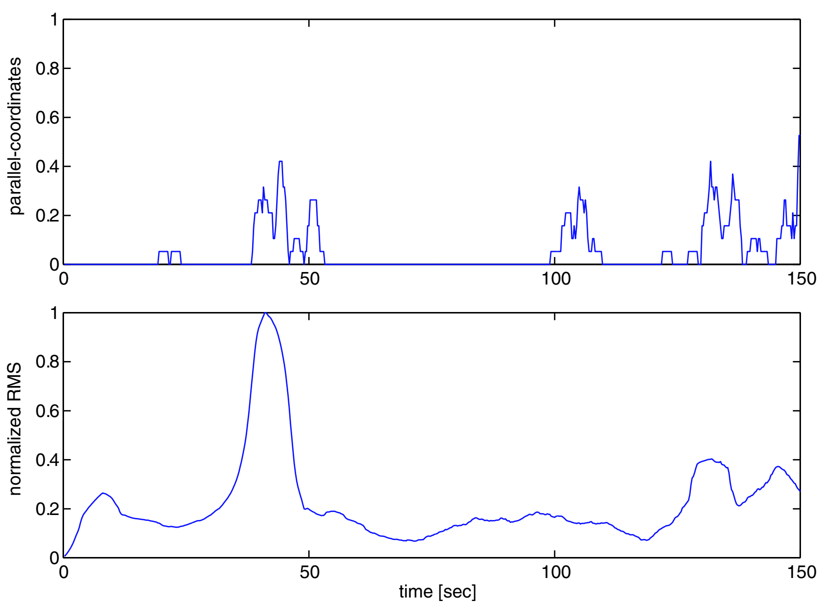

Fig. 17. Identification of a convoy of vehicles. All the vehicles in the convoy are of types $G^i, i \geq 5$. Note that none of the vehicles is falsely identified.

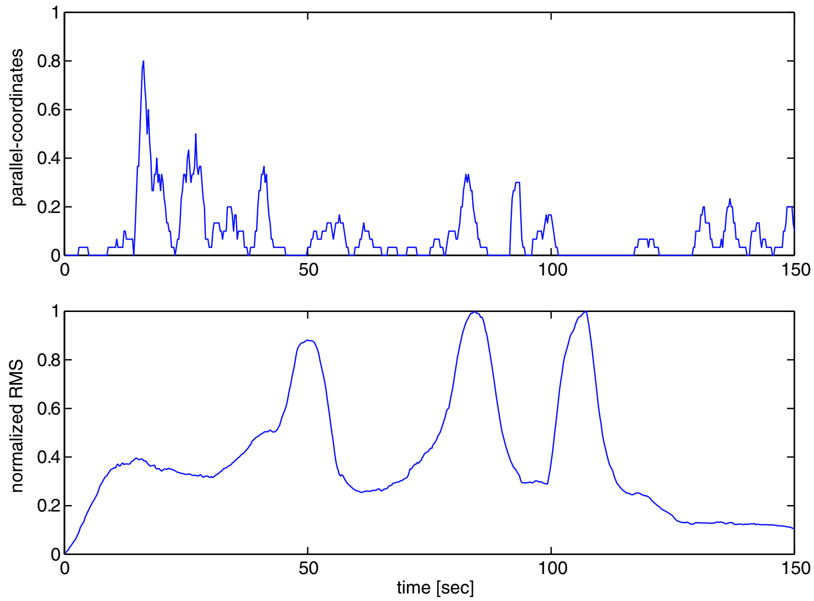

Fig. 18. Identification of a convoy of vehicles. All the vehicles are of types $G^i, i \geq 5$. Note that none of the vehicles is falsely identified.

identify the features (frequency bands) that enable one 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 four 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 on the base of local cosine transform in the frequency domain. The parallel coordinates analysis proved to be an efficient tool for defining a signature of a class of signals. 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. P. Auscher, 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* (1992), pp. 237–256.
2. A. Z. Averbuch, E. Hulata, V. A. Zheludev and I. Kozlov, A wavelet packet algorithm for classification and detection of moving vehicles, *Multidim. Syst. Sign. Process.* **12** (2001) 9–31.
3. L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone, *Classification and Regression Trees* (Chapman & Hall, 1993).
4. J. Buckheit and D. Donoho, Improved linear discrimination using time-frequency dictionaries, *Proc. SPIE* **2569** (1995) 540–551.
5. R. R. Coifman and Y. Meyer, Remarques sur l'analyse de Fourier a fenêtre, *C. R. Acad. Sci.* (1991) 259–261.
6. R. R. Coifman and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inf. Theory* **38** (1992) 713–719.
7. R. R. Coifman, private communication, 1998.
8. R. A. Fisher, The use of multiple measurements in taxonomic problems, *Ann. Eugenics* **7** (1936) 179–188.
9. A. Inselberg, Visual data mining with parallel coordinates, *Comput. Statist.* **13** (1988) 47–63.
10. A. Inselberg and B. Dimsdale, Multidimensional lines. 1. reperesentation, 2. proximity and applications, *SIAM J. Appl. Math.* **54** (1994) 559–577; 578–596.
11. A. Inselberg, *PARALLAX software package*, www.math.tau.ac.ilãiisreal.
12. R. R. Coifman, Y. Meyer and M. V. Wickerhauser, *Adapted waveform analysis, wavelet-packets, and applications*, in *Proc. of ICIAM'91* (SIAM Press, 1992), pp. 41–50.
13. Q. Jiang, S. S. Goh and Z. Lin, Local discriminant time-frequency atoms for signal classification, *Sign. Process.* **72** (1999) 47–52.
14. I. Kozlov, A. Averbuch and V. Zheludev, Wavelet packet based algorithm for identification of quasi-periodic signals, *Proc. SPIE* Vol. **4478**, *Wavelet Applications in Signal and Image Processing IX*, eds. A. Aldroubi, A. F. Laine and M. A. Unser (2001), pp. 353–360.
15. S. Mallat, *A Wavelet Tour on Signal Processing* (Acadamic Press, 1998).
16. S. Mallat and Z. Zhang, Matching pursuit with time-frequency dictionaries, *IEEE Trans. Sign. Procs.* **41** (1993) 3397–3415.
17. G. Matviyenko, Optimized local trigonometric bases, *Appl. Comput. Harmonic Anal.* **3** (1996) 301–323.
18. N. Saito, Local feature and its application using a library of bases, PhD thesis, Yale University, New Haven, 1994.
19. N. Saito and R. R. Coifman, Local discriminant bases and their application, *J. Math. Imaging Vision* **5** (1995) 337–358.
20. N. Saito and R. R. Coifman, Improved local discriminant bases using probability density estimation, *Proc. Amer. Statist. Assoc., Statist. Comput. Sec.* **5** (1996) 312–321.
21. W. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software* (AK Peters, 1994).