Wavelet based algorithm for acoustic detection of moving ground and airborne targets

Amir Z. Averbuch[a] Valery A. Zheludev [a] and Inna Kozlov[b]

[a]Department of Computer Science, Tel Aviv University
Tel Aviv 69978, Israel
[b] Electro-Optics Research and Development Ltd (EORD)
Technion, Haifa 32000, Israel

ABSTRACT

We detect the presence of a vehicle or an air borne target from a certain class via the analysis of its acoustic signature against an existing database of recorded and processed acoustic signals. To achieve this detection with no false alarms we construct the acoustic signatures of certain targets to be found by the distribution of the energies among blocks which consist of wavelet packet coefficients. We developed an efficient procedure for adaptive selection of the characteristic blocks. We modified the CART algorithm in order to utilize it as a decision unit in our scheme. A wide series of field experiments manifested a remarkable efficiency of the algorithm. The detection had been achieved practically with no false alarms even under severe conditions such as the acoustic recording of sought-after object was a superposition of the acoustics emitted from other targets that belong to other classes. The detection was even immune to severe noisy surroundings.

Keywords: Wavelet packets, discriminant blocks, acoustic signature, CART algorithm

1. INTRODUCTION

We present our results in solving an intrinsically interesting problem of detection of the presence of a vehicle or a flying aircraft belonging to a certain class via the analysis of the acoustic signals emitted while it is moving. The problem for the ground target is especially complex because of the great variability in the surrounding conditions of the recorded database: velocity, distances between the vehicles and the receiver, the presence of other vehicles, the roads the vehicles are traveling on, and the background noise, just to name few.

A crucial factor in having a successful detection (no false alarm) is to construct signatures built from characteristic features that enable to discriminate between the class of interest and the background. Multiscale wavelet analysis (see Ref. 10, Ref. 11) provides a promising methodology. However, existing wavelet based techniques (Ref. 12, Ref. 13, Ref. 5, for example) lack translation invariance in time domain. This is critical for detection, since misalignment between different signals can generate false results. Therefore, we developed a new technique 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. These blocks contain the distinctive characteristic features.

In the final phase of the process, in order to identify the acoustic signatures of the presented signals, we used properly modified conventional classifier Classification and Regression Tree (CART) Ref. 4.

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

Further author information: (Send correspondence to Amir Z. Averbuch)
Amir Z. Averbuch: E-mail: amir@math.tau.ac.il
Valery A. Zheludev : E-mail: zhel@math.tau.ac.il
Inna Kozlov: E-mail: eorddik@techunix.technion.ac.il

The paper is organized as follows. In Section 2 we formulate the approach to solve the problem. In Section 3 we describe the algorithm which is centered around two basic issues: Selection of the discriminant blocks of the wavelet packet transforms and discrimination of signals. Then, we explain its implementation. Section 4 presents the experimental results.

2. FORMULATION OF THE APPROACH

2.1. The structure of the signals

We assume that the recorded acoustic data generated by the vehicles belong to one of the two classes $C^k,\ k=0,1$. To Class C^1 we assign the signals recorded in presence of the vehicle to be detected and to Class C^0 we assign the signals recorded in absence of this vehicle. This class can include signals emitted by vehicles of other types and also the pure background. The recordings were taken during a wide range of field experiments under various surrounding conditions. In particular, the velocities of the vehicles and their distances from the recording device were varied. Moreover, the vehicles traveled on either paved (asphalt) or various unpaved roads, or on a mixture of paved and unpaved roads.

Figure 1 displays some portions of the signal emitted by a vehicle and its Fourier transform recorded at different conditions.

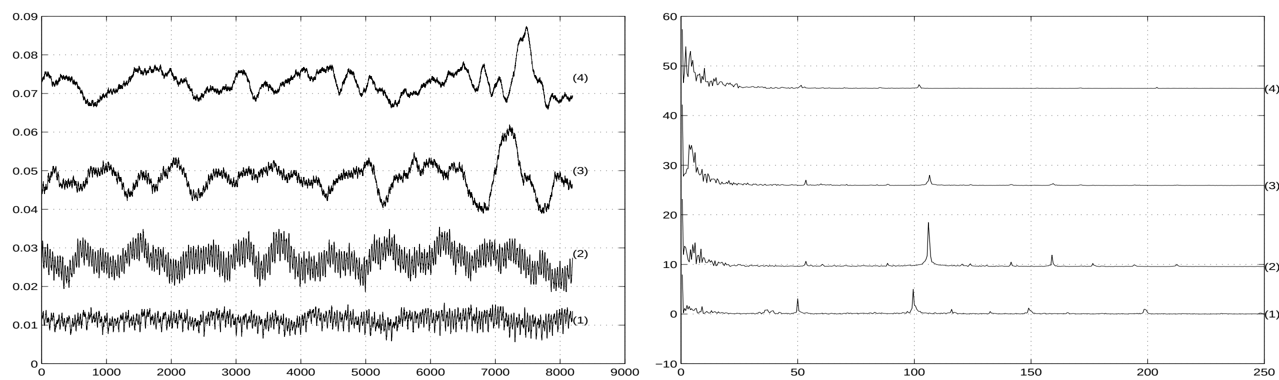

Figure 1. Left picture: sections of the signal recorded at different conditions: line 1–the vehicle is approaching the receiver; line 2–passing by the receiver; line 3–moving away from the receiver at a moderate speed; line 4–the same but far away from the receiver. Right picture: portions of the Fourier transforms of the signals shown in the left picture.

We realized that even within the same class the signals differ significantly from each other. The same is true for their Fourier transforms. However, there are some common properties to all the acoustic signals that were recorded from moving vehicles. First, these signals are quasi-periodic in the sense that there exist some dominating frequencies in each signal. These frequencies may vary with the change of the motion conditions. However, for the same vehicle these variations are confined in narrow frequency bands. Moreover, the relative locations of the frequency bands are stable (invariant) to some extent for signals that belong to the same class and differ for signals from other classes.

Therefore, we conjectured that the distribution of the energy (or some energy-like parameters) of signals that belong to some class over different areas of the frequency domain may provide a reliable characteristic signature for this class.

2.2. Formulation of the approach

Wavelet packet analysis is a highly relevant tool for adaptive search for valuable frequency bands of a signal or a class of signals. Once implemented, the wavelet packet transform of a signal yields a huge variety of different partitions of the frequency domain. One of the important features of the transform is its computational efficiency. The implementation of an m level transform requires $O(mn)$ operations. Due to the lack of time invariance in the multiscale wavelet packet decomposition, we will deal with the whole blocks of wavelet packet transform rather than with individual coefficients and waveforms. Moreover, following the suggestion in Ref. 8, we increase the number of sample signals 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.

General approach:

1. For training we use a set of signals with known membership. From them we select a few blocks which discriminate efficiently between the given classes of signals.
2. We apply the wavelet packet transform on the signal to be classified. We use as its characteristic features which are the normalized l_2 or l_1 norms of the wavelet packet coefficients contained in the selected blocks.
3. Finally, we submit the vectors of the extracted features to CART classifier. The latter, being appropriately trained beforehand, decides which class this signal belongs to.

3. DESCRIPTION OF THE ALGORITHM AND IMPLEMENTATION

3.1. The algorithm

The algorithm is centered around two basic issues:

I. Selection of the discriminant blocks of the wavelet packet coefficients is done according to the following steps:

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

II. Discrimination among the signals.

1. Preparation of the pattern set.
2. Building the CART classification trees.
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

3.2.1. Selection of discriminant blocks

Choice of analyzing waveforms: By now a broad variety of orthogonal and biorthogonal filters that generate wavelet packets coefficients are available. We use 8-th order spline wavelet packets and the Coiflet5 wavelet packets with 10 vanishing moments. These filters reduce the overlapping among frequency bands associated with different decomposition blocks. The empirical analysis suggests to decompose the signals into 5 to 7 levels (scales).

Construction of the training set: Initially, we gather as many recordings as possible for each class C^l, $l = 0, 1$ which have to be separated. Then we prepare from each selected recording, which belongs to a certain class, a number of overlapping slices of length $n = 2^J$ samples each, shifted with respect to each other by $s \ll n$ samples. All together we produce M^l slices for the class C^l. These groups of slices form the training set for the search of discriminant blocks.

Calculation of the energy map: First, we specified the kind of energy type measure to be used. Typically, we use either the normalized l_2 or l_1 norms of the blocks. After the measure has been chosen, the following operations are carried out:

1. The wavelet packet transform is applied up to scale m on each slice of length n from a given class C^l. This procedure produces mn coefficients arranged into $2^{m+1}-1$ blocks associated with different frequency bands.
2. The slice $A^l(i,:)$ is decomposed. The "energies" of each block are calculated in accordance with the chosen measure. As a result we obtain, to some extent, the distribution of the "energies" of the slice $A^l(i,:)$ over various frequency bands of widths from $N_F/2$ to N_F/m, where N_F is Nyquist frequency. It is presented by an energy vector E_i^l of length $2^{m+1}-1$.
3. The energy vectors along the training set of the class are averaged:

$$E^l = \frac{1}{M}\sum_{i=1}^{M} E_i^l.$$

The average energy map E^l of length $2^{m+1}-1$ indicates how the distribution of the "energies" among various block of the decomposition and frequency bands, respectively, is taking place within the whole class C^l. Similar operations are performed on both classes C^l, , $l=0$, 1.

The left picture in Figure 2 displays a typical energy map for 5 decomposition levels of a two-class problem. Heights of the bars indicate the normalized energy of each of the 63 decomposition blocks.

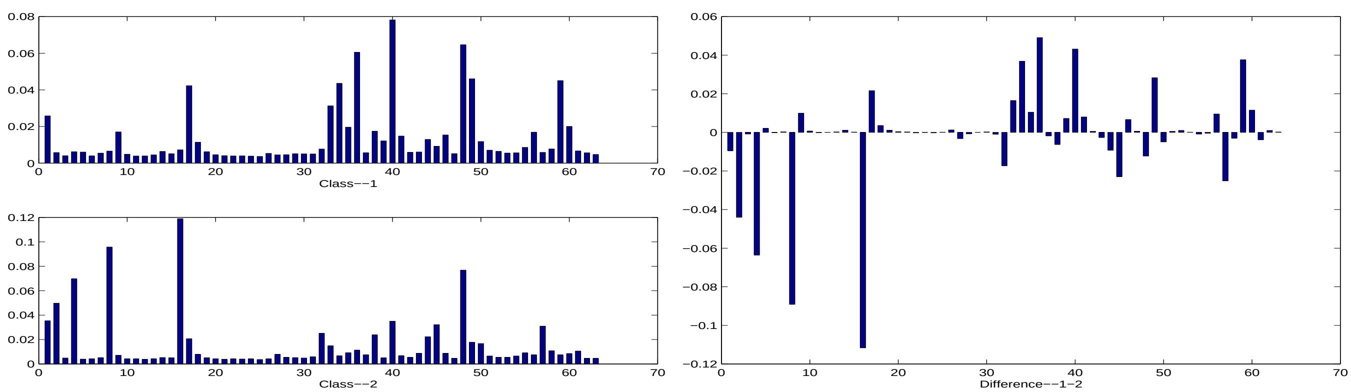

Figure 2. Energy map for 5 decomposition levels of a two-class problem (left picture) and difference of class C^1 and class C^2 maps (right picture). The length of a slice is $n = 1024$ samples.

Evaluation of the discriminant power of decomposition blocks and selection of discriminating blocks The average energy map E^l yields some sort of characterization for the class C^l but it is highly redundant. To gain a more concise and meaningful representation of the class we select the most discriminating blocks. One possible way to do so is the following. First, note that for a two-class problem the difference between two maps provides some insight into the problem (Figure 2 - right). As demonstrated in the figure, the differences for most blocks are near zero. It means that they are of no use for discrimination unlike a few blocks with large values in their differences. Therefore, the term-wise difference (absolute values) of the energy maps serves as the discriminant power map for the decomposition blocks $|E^1 - E^2|$.

Now we are in a position to select a few discriminant blocks which form a sort of signatures for the classes. This was not possible immediately since we are in a situation where the frequency bands of the blocks overlap. To avoid this frequency overlap we apply the procedure somewhat similar to the *Best Basis Selection Algorithm* Ref. 6, Ref. 12, Ref. 13. The idea is to compare the discriminant power of each pair of the "children" blocks with those of their parent blocks. In the case when the discriminant power of the parent exceeds sum of the children powers, the children blocks are discarded and vice versa. As a result we obtain some non-overlapping set of blocks which map the whole frequency domain of our signals, which are 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 the blocks with the highest discriminant factor. Moreover, if we are interested in certain frequency bands, we can select the corresponding blocks.

A slightly different approach is possible if we notice that blocks which gained positive difference values were stronger for the class C^1 and vice versa. Then we could select discriminant blocks separately from the positive and negative sub-maps of the difference map i.e. separately for each class.

Conclusion: As a result of the operations described above we discover a relatively small set of decomposition blocks such that the distribution of the energies amongst them characterize the classes to be distinguished. 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.

3.3. Classification

Once we have the set of discriminant blocks $B_1, ..., B_t$, we proceed to the classification phase.

Preparation of the reference set. Initially, we chose a number of recordings that belong to the classes C^l, $l = 0, 1$, to be distinguished, from which we form the reference set. These recordings are sliced similarly to that which was used for the preparation of the training set. For a certain class C^l, we form from each selected recording from the class a number of overlapping slices each of length n. These are shifted with respect to each other by s samples. We assume that there are μ^l slices related to the class C^l that are are gathered into $\mu^l \times n$ matrix a^l, $l = 0, 1$.

Then, we apply the wavelet packet transform up to level m on each row $a^l(i, :)$ of this matrix. After the decomposition of the slice $a^l(i, :)$, we calculate only the "energies" of the t blocks $B_1, ..., B_t$ that were selected before. In doing so we obtain the $1 \times t$ vector $V^l(i, :)$ which we regard as a representative of the slice $a^l(i, :)$. The vectors $V^l(i, :)$ form the $\mu^l \times t$ reference matrix V^l of the class C^l. We do the same for both classes C^l.

The reference sets are used as training sets for building CART. The construction of the tree is done by a binary split of the space of input patterns $X \longrightarrow \{X_1 \bigcup X_2 \bigcup \ldots \bigcup X_r\}$, so that, once a vector appeared in the subspace X_k, its membership could be predicted with a reasonable reliability. The basic idea with the split is that the data in each descendant subsets is more "pure" than the data in the parent subset. More details are given in the Appendix.

After the construction of the classification tree, we are in a position to classify test signals. To do so we must preprocess these signals.

Preparation of the test set. Suppose 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 (let it be K) of overlapping slices of length n each, shifted with respect to each other by s samples. All the K slices are gathered into $K \times n$ matrix T.

Each row $T(i, :)$ of this matrix is operated by the wavelet packet transform up to level m. In the decomposed slice $T(i, :)$, we calculate the "energies" of the t blocks $B_1, ..., B_t$ that were selected before. In doing so we obtain the $1 \times t$ vector $W(i, :)$ which we regard as a representative of the slice $T(i, :)$. The vectors $W(i, :)$ form the $K \times t$ test matrix W associated with the signal S.

Making the decision. Once the test matrix W is ready, we present each row $W(i, :)$ of the matrix to the CART classifier. CART uses the tree that was constructed before on the basis of the pattern sets. Once a vector is presented to the tree it is assigned to one of the subsets X_k of the input space X. This determines the most probable membership of the vector.

Then we count the numbers of vectors $W(i, :)$ attributed to each class C^l and make the decision in favor of the class which gets the majority of the vectors $W(i, :)$. The robustness of the decision is checked by the percentage of the vectors $W(i, :)$ attributed to this class.

4. THE RESULTS

We conducted two series of experiments to detect moving targets. In the first series the targets were the moving vehicles of a certain type and in the second series the targets were the flying aircrafts of a certain type. The acoustic sensors were placed on stationary platforms.

We obtain highly robust detection results for the ground targets. The results of airborne targets are preliminary but, nevertheless encouraging.

We tested various families of wavelet packets and various norms for the feature extraction and various combinations of features presented to the CART classifier. The best results were achieved with wavelet packets based on Coiflet5 filters with 10 vanishing moments and splines of order 8.

4.1. Detection experiments with the ground targets

Formally, our detection problem can be reduced to classification into two classes:

Class C^1 which contains signals emitted by the sought-for vehicles at various environments, velocities, distances and types of roads they traveled on.

Class C^0 which contains signals emitted by other types of different vehicles and pure background noises.

Such a diversity in the structure of Class C^0 signals could result in a high rate of false alarms. To reduce this rate during the selection process of discriminant blocks we choose blocks whose "energy" characterize Class C^1 signals. We preferred the CART classifier to others. Moreover, we modified the CART algorithm so that in order for a signal to be a member of Class C^1 it has to obey strict conditions that we imposed upon the classifier. These modifications, which are described in the Appendix, proved to be very efficient. The false alarm rate was reduced to almost zero, whereas by implementation of the standard CART it reached up to 64%. We observed that using hard-threshold denoising on the signals (see Ref. 9) to be classified improves the results. We had two groups of recordings of acoustic signals belonging to the classes $C^k,\ k = 0,\, 1$. The sampling rate was 4000 samples per second. Each group comprised a number of recordings of signals with length of 65536 samples (about 16-second).

We processed the signals using the scheme explained above. For the selection of the discriminant blocks we used 40 signals from each group recorded at various distances. The signals were processed under sliding overlapped windows of size $n = 1024$. The window was shifted along the signal with a step of $s = 128$ samples. Each window was decomposed by the wavelet packet transform into 6 or 7 levels (scales). The best results were achieved when we used either Spline8 or Coiflet5 wavelet packets and l_1 norm as the "energy" measure for the blocks. As a result of the procedures that were described in Section 3.2.1, we selected various sets of discriminant blocks.

The pattern set involved 10 signals from each group chosen from 40 signals which we used before for the selection of the discriminant blocks. The vectors of l_1 or l_2 norms of the selected wavelet packet blocks of the windows formed two reference matrices of size $5040 \times q$, where q is the number of the blocks in the given set. These matrices were used for the construction of the CART trees.

The algorithms were tested extensively in many experiments with no advance knowledge of the field conditions. The signals for the construction of the acoustic signature were the only ones that we know in advance.

We tested independently non-overlapping signals with duration of 2 seconds each that were retrieved from the recordings of classes C^1 and C^0. In other words, a signal was classified after 2-seconds of listening.

The results of some experiments are presented in figures 3-8.

The $*$ in the figures corresponds to a single 2-second signal of Class C^1. Let us take a signal S^l from the Class C^1. The corresponding $*\ s^l$ is located in position l which is related to the distance of the vehicle from the receiver. Its height, h^l, may range from 0 to 1 and is equal to $h^l = \frac{K_c^l}{K^l}$ where K^l is the total number of vectors $W(i,:)$ associated with the signal S^l, and K_c^l is the number of the vectors $W(i,:)$ attributed correctly to the class C^1. A 2-second signal is labeled as detectable if no less than 80% of its slices $W(i,:)$ belong to Class C^1. In all the figures the display is based on 2-second recordings (8000 samples). Moreover, under each diagram we depict the *loudness* level of each recording $R^n = \{r_k^n\}_{k=1}^{8000}$ expressed through root mean square $RMS = \sqrt{\sum_{k=1}^{8000} (r_k^n)^2}$.

Obviously, the *loudness* level corresponds to the distance of the vehicle from the sensor and it strongly depends on the atmospheric and background conditions. It is worth noting that the best results were achieved when we used the Coiflet wavelet packets with 10 vanishing moments and the signals were decomposed up to the 7th scale. In all the experiments presented below we used the same common set of discriminant blocks and the same classification tree (CART) was used in all the experiments.

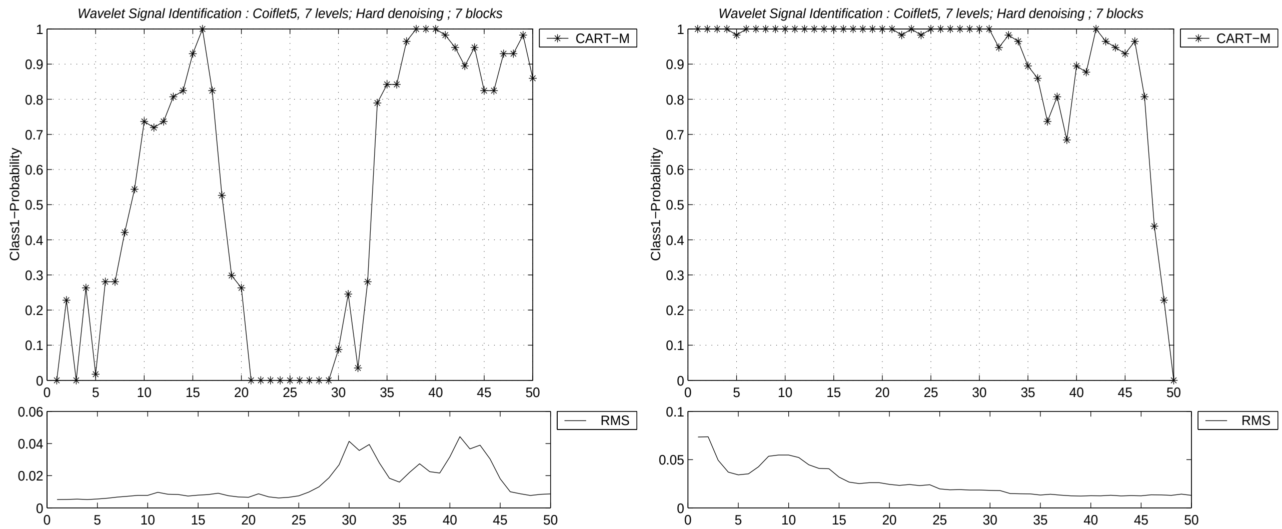

Figure 3. A vehicle from Class C^1 is moving away from the receiver up to a distance of 800 meters (left top picture) and from 800 meters to 1600 meters (right top picture) at 35 km/h. Bottom pictures depict the RMS of each 2-second signal. They all traveled on asphalt road. The following parameters were used: Coiflet5 wavelets, 7 blocks, modified CART classifier and every 2 seconds a decision was made.

In Fig. 3 we present the detection results of a single vehicle from Class C^1, which moves away from the receiver up to a distance of 1600 meters, at 35 km/h on an asphalt road. We can see that a satisfactory detection up to 1400 meters distance was achieved. It is interesting to note that at remote distances (> 300) meters the detection is even more stable than near the receiver. The drop in the reliability of the results in segments 23-28 in the left picture could be explained by a sudden change in atmospheric conditions. We note also that in this experiment the acoustic recording of the vehicle is not audible after 1100 meters.

In Fig. 4 we present an experiment similar to the previous one, with one difference: the vehicle approached the receiver from a distance of 1800 meters at 40 km/h. The initial distance was $\approx$ 1600 meters. The detection is remarkably stable.

The recordings In fig. 5 were taken while the vehicle traveled on the ground instead of traveling on an asphalt road. In the left pictures we present the results when the vehicle is moving away up to 1000 meters from the receiver. In the right picture the results were taken while the vehicle approached the receiver from 900 meters distance. The velocity again was 40 km/h. The range for having a valid acoustic recording was chosen to be 700 meters. One can see that our tool which detected successfully the target on the asphalt, did a good job on the ground road as well.

In the left pictures of Fig. 6 a more complicated situation is demonstrated. Here the vehicle from Class C^1 passed the receiver being accompanied by two vehicles from Class C^0 . The vehicles followed each another on an asphalt road at 40 km/h. One can see that, despite the strong unfavorable background noises, the vehicle from Class C^1 was satisfactorily detected.

In the right pictures of Fig. 6 we address the problem of misclassification (false alarms). While we used the standard CART algorithm, we were unable to address this problem well. It was resolved after a modification to the algorithm, as is explained in Appendix, was made. The results of the experiment displayed in these pictures are similar to the previous one with the difference that the vehicle from Class C^1 is missing from the convoy. The plot with the circles depicts the results of classification with 2-second fragments by the conventional CART algorithm. One can observe that 8 fragments erroneously detected a Class C^1 vehicle with probability exceeding 80%. This means that a definite false alarm occurred. When the modified CART was used and this is depicted by the plot with asterisks, the classification results which indicate that it does not belong to Class C^1, was with probability exceeding 37%. In the other experiments we also had a very low misclassification rate while using the modified CART.

The following is a hard problem that remains unsolved. The goal is to detect vehicles of Class C^1 moving within a large convoy of other vehicles. The detection must be fulfilled even in the event when some vehicles from Class C^0

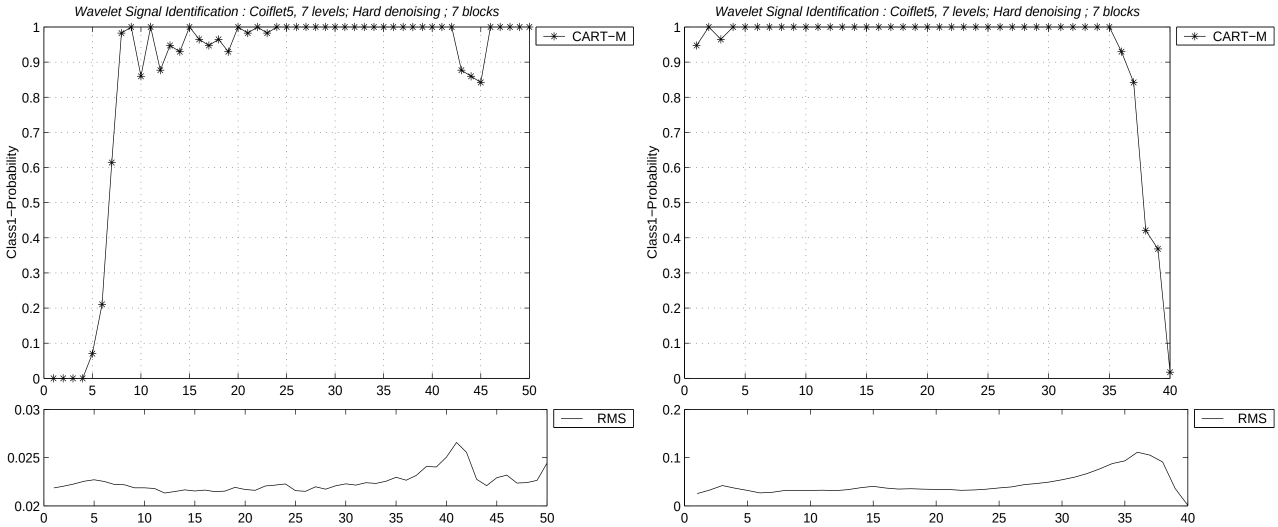

Figure 4. A vehicle from Class C^1 is approaching the receiver from a a distance of 1800 meters till it reaches 900 meters (left top picture) and from 900 meters (right top picture) at 40 km/h. Bottom pictures depict the RMS of each 2-second signal. They all traveled on asphalt road. The following parameters were used: Coiflet5 wavelets, 7 blocks, modified CART classifier and every 2 seconds a decision was made.

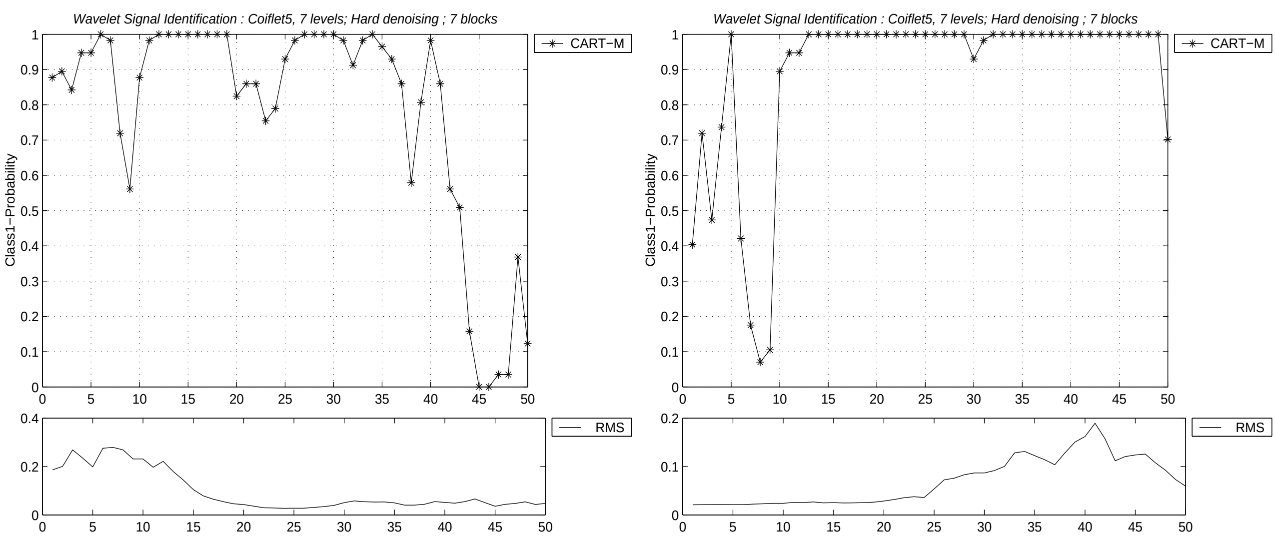

Figure 5. A vehicle from Class C^1 is moving away from the receiver up a a distance of 1000 meters (left top picture). The vehicle approached the receiver from 900 meters distance at 40 km/h. Bottom pictures depict the RMS of each 2-second signal. They all traveled on ground road. The following parameters were used: Coiflet5 wavelets, 7 blocks, modified CART classifier and every 2 seconds a decision was made.

shadow the sought-for targets passing near the receiver. These situations are illustrated in Fig. 7. In the left picture the convoy consists of 10 vehicles and two of them belong to Class C^1 . Fragments $20-40$ are related to the moment when two vehicles from Class C^0 passed near the receiver shadowing the vehicles from Class C^1 to be detected. We can see that for these fragments the detection rate is plunging to near zero. The velocity here was 28 km/h. In the right picture the convoy consists of 6 vehicles and two of them belong to Class C^1 . It approaches the receiver at a rate of 56 km/h. Fragments $10-30$ are related to the time interval when two vehicles from Class C^0 shadowed the vehicles from Class C^1 to be detected. It is worth noting that without the presence of this shadowing we sometimes

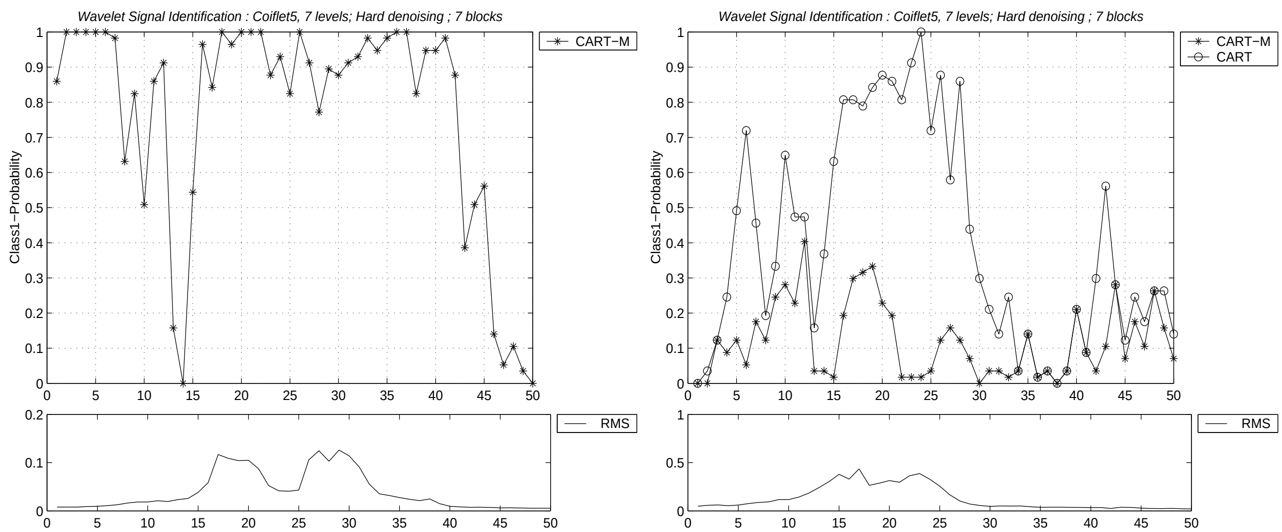

Figure 6. Left pictures: A vehicle from Class C^1 is passing the receiver accompanied by two vehicles from Class C^0. Right pictures: Two vehicles from Class C^0 without the vehicle from Class C^1 are passing the receiver. The plot with circles depicts results of classification of 2-second fragments by the conventional CART algorithm. The plot with asterisks does the same for the modified CART. All is done while the vehicles travel at 40 km/h on asphalt road. Bottom pictures depict the RMS of each 2-second signal. The following parameters were used: Coiflet5 wavelets, 7 blocks. Decision was made every 2 seconds.

get a satisfactory detection.

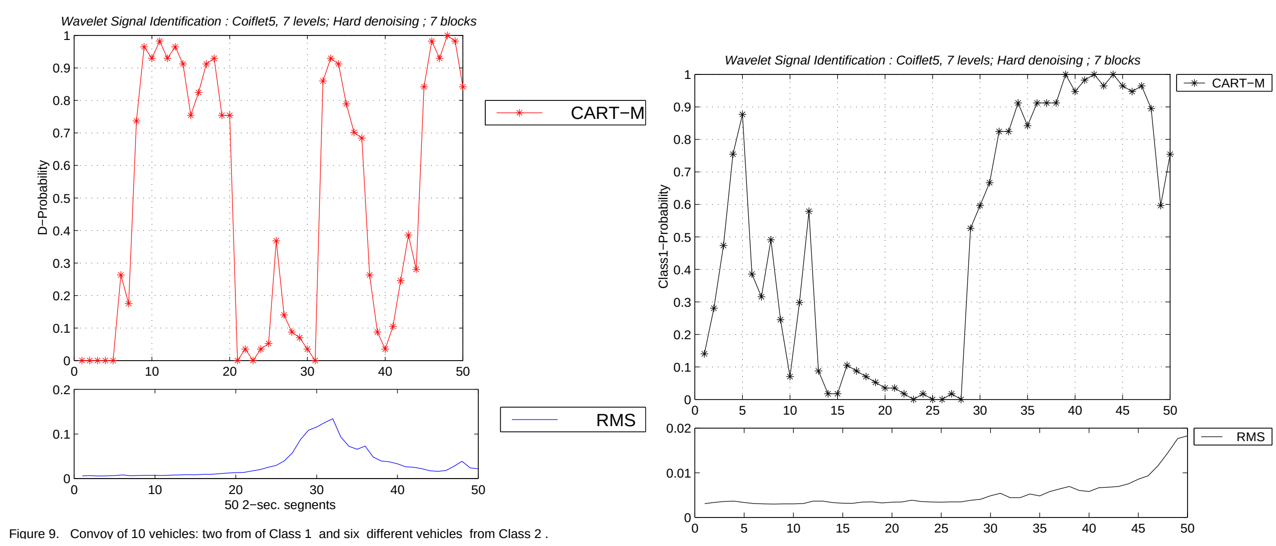

Figure 7. Left pictures: The convoy contains 10 vehicles including two vehicles which belong to Class C^1. Fragments $20-40$ are related to the time interval when two vehicles from Class C^0 passed near the receiver shadowing the vehicles of Class C^1 to be detected. The velocity is 28 km/h. Right pictures: The convoy contains 6 vehicles and two of them belong to Class C^1 . It approaches the receiver at a rate of 56 km/h. Fragments $10-30$ are related to the moment when two vehicles from Class C^0 shadow the vehicles from Class C^1 to be detected.

4.2. Airborne targets

We conducted a number of experiments to detect flying aircrafts of a certain type. We used the same algorithms that were used for ground targets. The only difference was that the decision was made after 1 second of listening. The first series of experiments gave a positive confirmation to the relevance of this approach. In Figure 8 we display results of two detection experiments. In the current experiments only short recordings were available this is in contrast to the ground experiments.

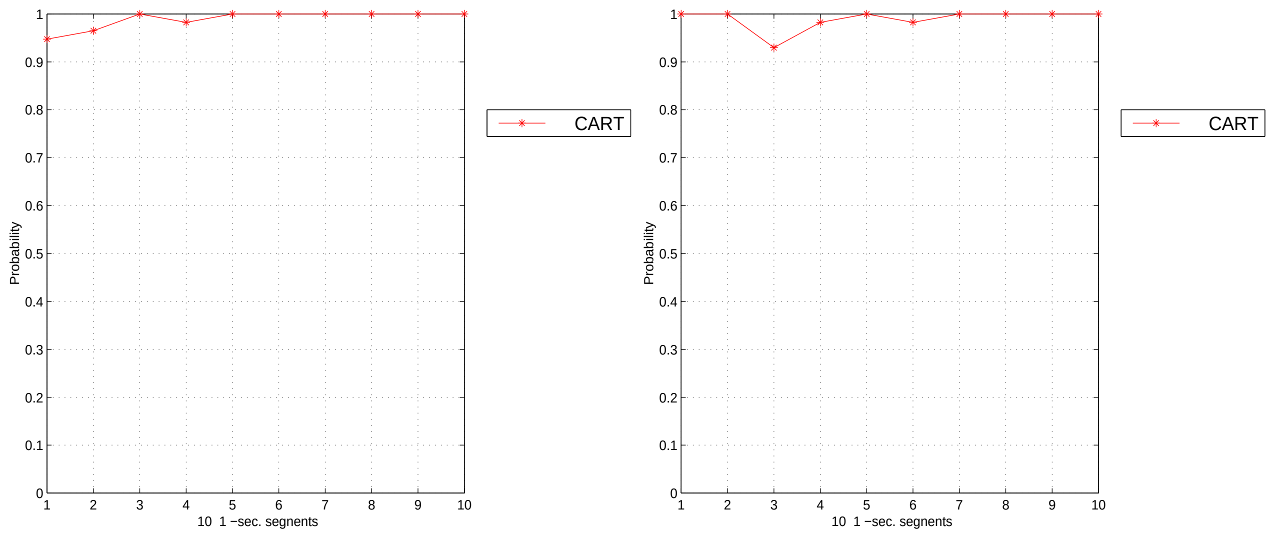

Figure 8. Results of detection of a flying aircraft of a certain type. Each ∗ corresponds to identification of a 1-second signal. The following parameters were used: Coiflet5 wavelets, 7 blocks, modified CART classifier and every 1 second a decision was made.

5. CONCLUSIONS

We presented a robust algorithm that solves the problem of detection of the presence of a vehicle or an aircraft from a certain class via the analysis of its acoustic signature against an existing database of recorded and processed acoustic signals.

The problem is important because it is applicable to real industrial problems. To achieve this detection with minimal false alarms, we constructed the acoustic signature of a target using the distribution of the energies among blocks which consist of wavelet packet coefficients. As decision unit we used Classification and Regression Tree (CART) classifier. We developed an efficient procedure for the selection of the most discriminating blocks of wavelet packet coefficients. Moreover, we enhanced the detection abilities of the classifier by a proper modification. By utilizing this we succeeded in reducing the false alarm rate to almost zero for majority of typical situations. The targets were reliably detected even at large distances from the receiver when the signals were completely inaudible to the human ear. We got satisfactory detection results for ground targets moving within a convoy of other vehicles.

But this investigation highlighted some unsolved problems. We want to achieve detection with a very unfavorable signal-to-noise ratio (SNR). Such situations arise, for example, when the target recedes at a large distance or when it is shadowed by other vehicles passing near the receiver or the acoustic sensor is placed on a moving platform. To tackle this problem, we intend, in further investigation, to devise efficient denoising procedures and to refine the selection of characteristic features of the sought-after signals. For this purpose we will employ the recently developed library of biorthogonal wavelets with strong adaptation abilities Ref. 2, Ref. 3.

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

6. APPENDIX: CART ALGORITHM AND ITS MODIFICATION

6.1. Outline of the standard scheme

A comprehensive exposition of the CART scheme can be found in Ref. 4. For simplicity of the presentation we consider a two-class classification problem.

Construction the tree

The space X of input patterns from the reference set consists of two reference matrices $V^l,\ l = 1, 2$ of sizes $\mu_l \times n$, respectively. We assume that $\mu_1 = \mu_2$. The i -th row of the matrix V^l is a vector $V^l(i, :)$ of length n representing the signal s_i^l which belongs to the class C^l. In our case, n is equal to the number of discriminant blocks. All row vectors $V^l(i, :)$ should be normalized, i.e.

$$0 \leq V^l(i, j) \leq 1, \qquad \sum_{j=1}^{t} V^l(i, j) = 1.$$

The tree structured classifier to be constructed has to divide our space X into J disjoint subspaces

$$X = \bigcup_{\nu=1}^{J} X_\nu^t. \tag{6.1}$$

Each subspace X_ν^t must be "pure" in the sense that the percentage of vectors from one of the matrices V^l, must prevail the percentage of the vectors from the other matrix. (In the original space X both are 50%.)

The construction of the binary tree is started by a split of X into two descendant subspaces:

$$X = X_1 = X_2 \bigcup X_3, \quad X_2 \bigcap X_3 = \emptyset.$$

To do so, CART chooses a split variable y_j and split value z_j in a way to achieve minimal possible "impurity" of the subspaces X_2 and X_3. The split rule for the space X_1 is as follows :

If a vector $y = (y_1, ...y_n)$ satisfies the condition $y_j \leq z_j$, then it is transferred to the subspace X_2 , otherwise it is transferred to the subspace X_3. In addition, we divide the subspace X_2 in a similar manner:

$$X_2 = X_4 \bigcup X_5, \quad X4 \bigcap X_5 = \emptyset.$$

The subsequent split variable y_k and split value z_k are selected so that the data in each of the descendant subspaces were "purer" than the data in the parent subspace. Then, one of the subspaces X_4 or X_5 can be further divided recursively until we reach to the so called terminal subspace X_1^t which is not splitted further. The decisions whether a subspace is classified as terminal subspace depends on the predetermined minimal "impurity" and the minimal size of the subspace. The terminal subspace X_1^t is assigned to the class C^l, with probability

$$p_1^l = \frac{m_1^l}{m_1} = \frac{\#\{y \in X_1^t \bigcap V^l\}}{\#\{y \in X_1^t\}},$$

where m_1^l is the number of points in node X_1^t that belongs to class C^l and m_i is the total number of points in the subspace X_1^t. After termination is reached in the subspace X_1^t we return to subspace X_3 which was not splitted. Similarly, we reach the next terminal subspace X_2^t. We do the same with one of yet non-splitted subspaces and finally the tree of (6.1) is constructed. In the terminology of graph theory, the space X is called the root node, the nonterminal and terminal subspaces are the nonterminal and terminal nodes.

Classification

A vector $x = (x_1, ...x_n)$ is fed to the tree in the first step is assigned to either node X_2 in the case where the coordinate $x_j \leq z_j$ or to node X_3 otherwise. Finally, by checking subsequent split variables, the vector is forwarded into a terminal node X_r^t which is labeled as class C^l, with probability p_r^l.

6.2. CART Modification

The main difference between our specific detection problem and the conventional two-class classification problem lies in the fact that the second (background) class is not properly defined and may contain a huge diversity of signals. While using the standard CART algorithm, we got an unacceptable false alarm rate (FA). For some vehicles from Class C^2, FA ranged up to 64%. To overcome this drawback, we modified the CART algorithm in order to have a detection tool which is based on it. The algorithm has the following steps:

1. Given the space X of input patterns, we build the classification tree as it was explained above (section 6.1). Thus, the space X is split into a set of disjoint terminal subspaces (nodes) $\{X^t_\nu\}$ – (see Eq. 6.1).
2. The probability value P is predetermined. Usually, the value of P is set to 0.75. We select from the set $\{X^t_\nu\}$ only the terminal nodes which are assigned to class C^1 with a probability exceeding P. Lets denote such nodes as $\{\Xi^j\}_{j=1}^L$.
3. Let Ξ^j be one of such terminal nodes. Denote $Y^j = V^1 \bigcap \Xi^j$. It means that this is a subset of all vectors $\{y^s = (y^s_1, ...y^s_n)\}_{s=1}^{S^j}$ from the matrix V^1 within the node Ξ^j.
4. We compute
$$\underline{y^j_i} = \min_s y^s_i, \quad \overline{y^j_i} = \max_s y^s_i, \; s = 1, \ldots, S^j.$$
Usually, we do it for all the coordinates $i = 1, \ldots, n$. But sometimes it is reasonable to handle only the most significant coordinates. Similar computation is performed with all the other terminal nodes $\{\Xi^j\}_{j=1}^L$.
5. Let us define the clusters $\{\Gamma^j\}_{j=1}^L$ as follows. We say that a vector $x = (x_1, ...x_n)$ belongs to a cluster Γ^j if all (or significant) coordinates satisfy the conditions
$$\underline{y^j_i} \le x^s_i \le \overline{y^j_i}, \quad i = 1, \ldots, n. \tag{6.2}$$
6. **Decision rule:** We assign a vector $x = (x_1, ...x_n)$ to Class C^1 if it belongs to one of the clusters Γ^j.

This modified CART algorithm reduced the FA in our experiments almost to zero.

REFERENCES

1. A. Averbuch, F. G. Meyer, J.-O. Strömberg, Fast adaptive wavelet packet image compression, *IEEE Trans. on Image Processing*, Vol. 9, No. 5, pp. 792-800, May 2000.
2. A. Averbuch, V. Zheludev, "Construction of biorthogonal discrete wavelet transforms using interpolatory splines" , *Tel Aviv University, The School of Mathematical Sciences, Technical Report,* **337**, (1999).
3. A. Averbuch, V. Zheludev, A. Pevnyi, "A lifting scheme of biorthogonal wavelet transform based on discrete interpolatory splines", This volume.
4. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees,* Chapman & Hall, Inc., New York, 1993.
5. J. Buckheit, D. Donoho, "Improved linear discrimination using time-frequency dictionaries",*Proc. SPIE,* **2569**, (1995), 540-551.
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, Y. Meyer, M. V. Wickerhauser, "Adapted waveform analysis, wavelet-packets, and applications", *In Proceedings of ICIAM'91,* SIAM Press, Philadelphia, 1992, 41-50.
8. R. R. Coifman, private communication, (1998).
9. 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.
10. I. Daubechies, *Ten lectures on wavelets.* SIAM, 1992.
11. S. Mallat, *A wavelet tour of signal processing,* Acad Press, 1998.
12. N. Saito and R. R. Coifman, "Local discriminant bases and their application", *J. Mathematical Imaging and Vision,* **5,** (1995), 337-358.
13. N. Saito and R. R. Coifman, "Improved local discriminant bases using probability density estimation", *Proc. Am. Statist. Assoc., Statistical Computing Section,* **5,** (1996), 312-321.