

RESEARCH PAPER

Estimation of efficiency of a tree structured hierarchical wavelet representation of synthetic database applied to non-cooperative target recognition

CHRISTIAN BROUSSEAU

In this paper, problem of efficient representation of large database of target radar cross section (RCS) is investigated in order to minimize memory requirements and recognition search time, using a tree structured hierarchical wavelet representation. Synthetic RCS of large aircrafts, in the High Frequency (HF)–Very High Frequency (VHF) bands, are used as experimental data. Hierarchical trees are built using wavelet multiresolution representation and K-means clustering algorithm. Criteria used to define these hierarchical trees are described and the obtained performances are presented.

Keywords: Radar signal processing and system modeling, Radar applications

Received 14 October 2010; Revised 12 April 2011; first published online 18 May 2011

I. INTRODUCTION

Requirements for future air defence radar systems are detection, localization, but also identification of aircrafts. With the increasing resolution of modern radar systems, it is theoretically possible to store much information, like aspect, elevation, pulse width, etc., of a complex target and use them in the field of target identification.

The advantage of increasing resolution of radar systems is the opportunity to have more characteristic details of a specific target. The disadvantage is that these detailed characteristics require much computer memory to be stored, computer resources, and increase the search time to non-cooperative target recognition (NCTR) association. It is therefore important to develop efficient methods to reduce the size of high-resolution data of radar targets. One way to compress these data is to use tree structured representation based on clustering algorithm coupled with a multiresolution wavelet representation to reduce the data size and the number of radar cross section (RCS) signature [1].

In this paper, we investigate the problem of efficient representation of large database of radar range profiles in order to minimize memory requirements and recognition search time, using a tree structured hierarchical wavelet representation.

The paper is organized as follows. In a first step, the used synthetic RCS database of large aircrafts is described. In a second step, after a brief review of wavelet transform theory

and unsupervised clustering theory, the method of building hierarchical trees is presented. Finally, in a third step, efficiency of an example of tree structured hierarchical wavelet representation using wavelet transform and K-means clustering algorithm are discussed regarding some criteria, like probability of false classification (P_{fc}) as a function of signal-to-noise ratio (SNR), minimum SNR to obtain a P_{fc} smaller than 1% and search computational time (S_{ct}) for a fixed SNR. Then, a comparison to other data compression techniques is performed.

II. DESCRIPTION OF THE SYNTHETIC RCS DATABASE

A) Introduction

The synthetic RCS database has been developed during the MOSAR project [2, 3] with the support of the French Ministry of Defence (DGA – Direction Générale de l'Armement).

The objectives of MOSAR project are to improve knowledge of frequency response of targets in resonance region by measurements, and to test efficiency of recognition methods. These studies led to:

- Development of a coherent, pulsed, quasi-monostatic, multifrequency, HF–VHF radar using the 20–100 MHz frequency band and horizontal and vertical polarizations.
- Development and validation of a simulated RCS database using numerical models of aircrafts in the 20–100 MHz frequency band.
- Development and tests of NCTR algorithms.

IETR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France.
Phone: +33 2 2323 6231
Corresponding author:
C. Brousseau
Email: christian.brousseau@univ-rennes1.fr

B) Description of the synthetic RCS database

To study aircraft RCS, several possibilities exist. One can perform:

- Anechoic chamber measurements on real aircrafts or scaled models.
- In-flight measurements with a radar system.
- Simulations using a computational model.

Anechoic chamber measurements are not well suited to collect data at various angles of a target but they are useful to validate numerical models. To perform in-flight measurements, it is necessary to use a calibrated radar system and to wipe out propagation effects. The simulation of RCS behavior, using a computational model, is a very attractive scheme but the model must be validated.

To be able to use a small computer like a PC, the simulation of RCS has been made with the free Numerical Electromagnetic Code NEC2 that is based on the method of moments [4, 5]. In this case, the aircraft structure is considered as a perfect electric conducting body. An example of wiregrid model is presented in Fig. 1. Currently, the synthetic database is constituted of eight mid-range airplanes: Airbus A320, BAe 146-200, Boeing 727-200, 737-200, 737-300, 747-200, 757-200, and Fokker 100.

For each aircraft, RCS has been determined using the following parameters:

- Frequency band: 20–100 MHz with a 1 MHz frequency step.
- Azimuth angle: -10 to $+190^\circ$ with a step of 2° .
- Elevation angle: $0-90^\circ$ with a step of 1° .
- Polarization: HH, HV, VH, VV.

The range profile is estimated from the 61 samples of the frequency response using an inverse fast Fourier transform with a zero padding on 256 points. So, the range step is 0.6 m. Then, a Hamming windowing is applied on the estimated range profile. The synthetic database is finally constituted of around 300 000 range profiles. Figure 2 shows an example of estimated range profile.

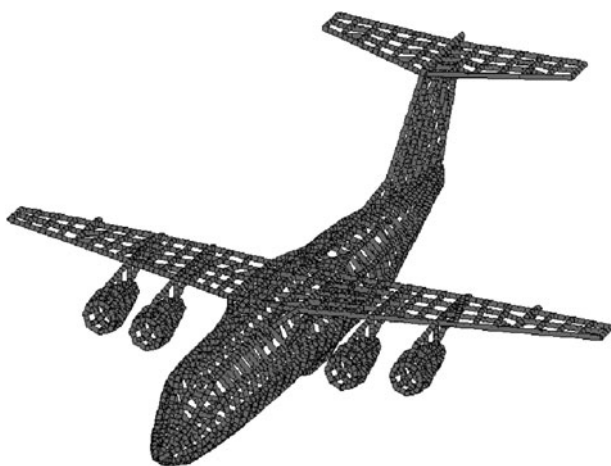


Fig. 1. Example of aircraft modeling using a wiregrid model – BAe 146-200.

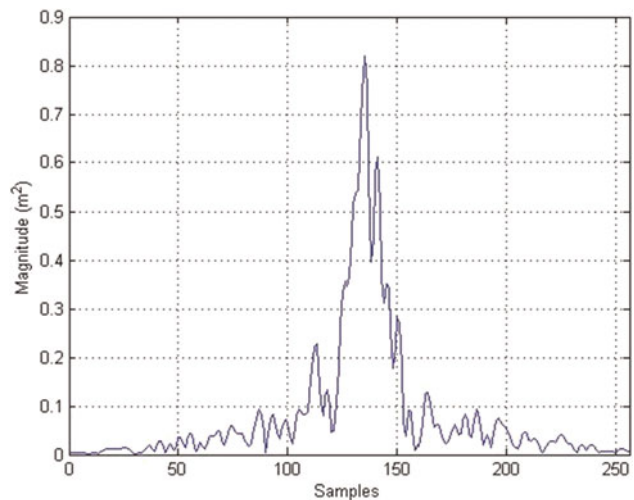


Fig. 2. Example of estimated range profile – BAe 146-200 – HH polarization – frequency band: 20–100 MHz.

III. APPLICATION OF TREE STRUCTURED HIERARCHICAL WAVELET REPRESENTATION TO DATABASE COMPRESSION

A) Introduction

Wavelet transforms and clustering algorithms are useful in a variety of applications. Wavelets provide the analyst with an approximation of the signal and a detail of the signal as well. Clustering aims at finding a structure in a collection of unlabeled data.

These methods have powerful efficiencies. Nevertheless, each of them has its own limitations. In this case, it should be interesting to merge them to build multiresolution hierarchical trees.

For a complete description of wavelet analysis and clustering algorithms, the reader should refer to [6–14]. A brief summary of how wavelets and clustering were used is presented below.

B) Multiresolution wavelet representation of RCS database

The discrete wavelet transform of finite sequences analyzes a signal S by decomposing it into approximations A_i and details D_i by quadrature filter systems [6, 7], where i is the decomposition level.

The approximations and details are, respectively, obtained by a low-pass filter and a high-pass filter. At each level, filtering process is followed by decimation by 2 that decreases the data size. Figure 3 presents the scheme of the filter system. Finally, the approximation and detail signatures at each level are pre-processed from the original signal and placed in the training data set. Figure 4 plots an example of range profile and its wavelet decomposition computed in four levels using the Haar mother wavelet.

A previous work [15] has shown that there is no statistically significant difference in performance of the classifier when different wavelets are chosen. Thus, in the next sections, only results obtained with the Haar wavelet are presented.

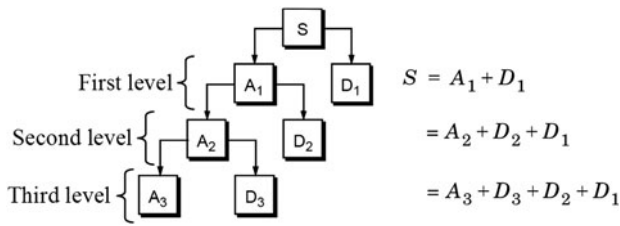


Fig. 3. Filter system of the wavelet transform (where S is the original signal, A_i the approximations, D_i the details, and i the decomposition level).

C) Unsupervised clustering of RCS database

Clustering can be considered as the most important unsupervised learning problem. It deals with finding a structure in a collection of unlabeled data. Clustering is the classification of objects according to similarities among them, and organizing of data in groups.

A popular measure to determine this similarity is the Minkowski metric [10]:

$$m(x_i, x_j) = \left(\sum_{k=1}^{dim} |x_{i,k} - x_{j,k}|^P \right)^{1/P}, \quad (1)$$

where dim is the dimensionality of the data, and $P \geq 1$ is a control of the distance growth of patterns. In our case, we have chosen to use the Euclidean distance where $P = 2$.

Two types of clustering methods can be defined [9]:

- Hard clustering techniques where data are set into C specified number of mutually exclusive subsets.

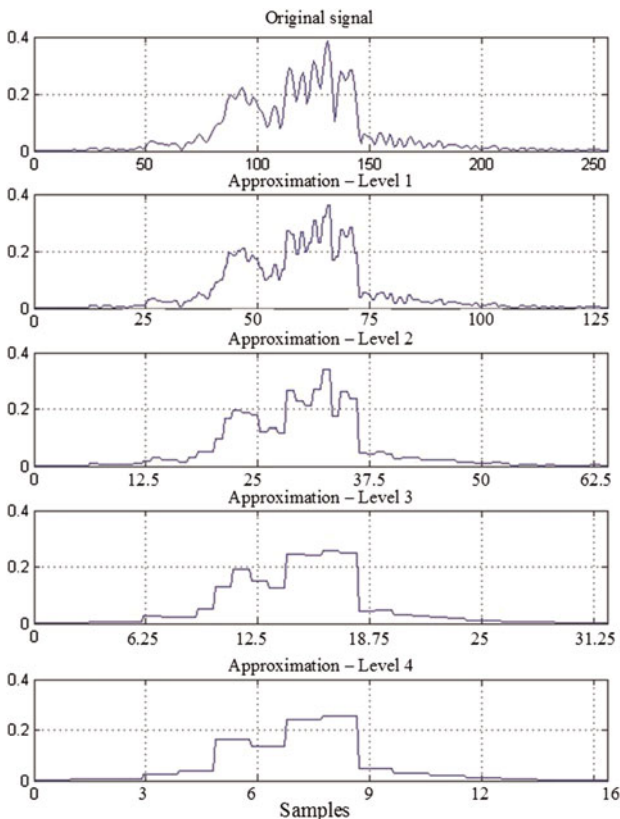


Fig. 4. Example of range profile and its wavelet transform computed in four levels using the Haar mother wavelet.

- Fuzzy clustering techniques where data can be assigned to several clusters simultaneously, with different degrees of membership.

Data membership to a partition is usually defined by an appropriate matrix U whose factors are equal to 0 or 1 in the case of hard clustering method, or a number between 0 and 1 in the case of fuzzy clustering method. For a complete description of these unsupervised clustering algorithms, the reader should refer to [8–14].

Previous results [15] have shown that better performances are obtained with a hard clustering algorithm, like K -means, in NCTR applications. Thus, in the next sections, only results obtained with the K -means hard partitioning method are presented.

The K -means hard partitioning method is simple and popular [11]. From an $N \times n$ dimensional data set, K -means algorithm allocates each data point to one of C clusters to minimize the following objective function:

$$J = \sum_{i=1}^C \sum_{k \in A_i} \|x_k - c_i\|^2, \quad (2)$$

where A_i is a set of objects (data points) in the i th cluster and c_i is the mean for those points over cluster i .

Thus, c_i are called the cluster centers (centroids) and are defined as

$$c_i = \frac{\sum_{k=1}^{N_i} x_k}{N_i}, \quad x_k \in A_i, \quad (3)$$

where N_i is the number of objects in A_i .

D) Multiresolution hierarchical tree representation of RCS database

These previous techniques are very useful in many applications. These methods give powerful efficiencies but each of them has its own limitations [15, 16].

Application of wavelets representation to NCTR application slightly decreases recognition search time but with a low degradation of false identification probability. At the opposite, use of clustering algorithm gives a very low decrease of recognition search time but with an important degradation of false identification probability. A way to improve these techniques is their association in a multiresolution hierarchical tree.

IV. TREE STRUCTURE DESIGN

In the case of clustering algorithm applied to NCTR, best efficiencies are obtained for an optimum number of clusters [16]. For a multiresolution hierarchical tree, problem is quite different. Number of clusters must be large to decrease computational time, but false identification probability must not be degraded.

Number of clusters on each decomposition level can be defined as a function of the distortion on the entire population of data vectors [1]. This distortion can be determined using a mean squared distance metric and is computed using the finest representation of the data vectors. It is defined as

$$d_{(j,c,o)} = \frac{n_c}{M} \sum_{i=1}^{n_c} \|C_{(j,c,o)} - S_i^o\|^2, \quad (4)$$

where j is the decomposition level, n_c the number of data vectors in cluster c , $C_{(j,c,o)}$ the centroid of the cluster re-sampled at the finest resolution (o), S_i^o the i th data vector at resolution o , and M the number of all data vectors.

For the entire tree, total distortion TD can be computed as

$$TD = \sum_{j=1}^J \sum_{c=1}^k d_{(j,c,o)}, \quad (5)$$

where J is the maximum decomposition level and k the number of clusters.

Then, to design the multiresolution hierarchical tree, processing steps are the following:

- *Step 1:* Loading of complete target database.
- *Step 2:* Wavelet decomposition of target database on the different levels.
- *Step 3:* Computation of clustering database on the lowest (coarsest) decomposition level using the $d_{(j,c,o)}$ criterion.
- *Step 4:* Computation of clustering database using the next finer resolution based on the previous subpartition and the $d_{(j,c,o)}$ distortion criterion.
- Repeat step 4 until the decomposition level o corresponding to the finest resolution (original signals).

Once tree is built, a pruning is realized by inspecting the contents of the different clusters. If a cluster contains only signatures of one aircraft or if on the upper level, the node has no leaves, then the branch is pruned.

To evaluate the consistency of the hierarchical tree, the total distortion TD and the entropy of the final partition E can be determined as a function of the number of clusters. Entropy is a measure of randomness of the population of a cluster and is defined as

$$E = - \sum_{j=1}^J \sum_{c=1}^k \frac{n_c}{M} \log\left(\frac{n_c}{M}\right). \quad (6)$$

V. PERFORMANCE ESTIMATION METHOD

A) Introduction

To test efficiency of database compression using multiresolution hierarchical trees, many criteria can be used:

- Probability of false classification (Pfc) as a function of signal-to-noise ratio (SNR).
- Minimum SNR to obtain a Pfc smaller than 1%.
- Search computational time (Sct) for a fixed SNR .

B) Probability of false classification

Probability of false classification Pfc is defined for M target classes as

$$Pfc = \frac{1}{M} \sum_{i=1}^M \frac{m_i}{n_i}, \quad (7)$$

where m_i is the number of classification error and n_i the number of element in the class i .

The nearest-neighbor algorithm [17] is used to recognize the target. It is a simple algorithm and is useful to test the efficiency of database compression using multiresolution hierarchical trees. The distance used to find the nearest neighbor is the Euclidean distance $d_T^{k,r,s}$ between the RCS magnitudes A_j :

$$d_T^{k,r,s} = \sqrt{\sum_{j=1}^N |A_j^{k,r,s} - A_j^T|^2}, \quad (8)$$

where j is the sample index, N the length of the signal, A_j^T the magnitude of measured RCS, and $A_j^{k,r,s}$ a database element (aircraft k , azimuth angle r , elevation angle s).

Then, minimal distances to each aircraft are computed and the nearest neighbor k_T for the measure T is extracted like

$$k_T = \arg \min_k (d_T^{k,r,s}). \quad (9)$$

C) Signal-to-noise ratio

To see the effect of random noise, zero-mean white Gaussian noise has been added to the signal. The signal-to-noise ratio (SNR) is defined as

$$SNR \text{ (dB)} = 10 \log_{10} \frac{\sum_{j=1}^N s_j^2}{N\sigma^2}, \quad (10)$$

where s_j is the sampled signal, N the length of the signal, and σ^2 the variance of Gaussian noise.

D) Search computational time

In computing, to estimate the “search computational time” (Sct), a standard parameter is the number of MFLOPs (Million Floating point Operations) with which comparison between efficiencies of different processing algorithms can be easily performed.

VI. APPLICATION OF MULTIREOLUTION HIERARCHICAL TREE TO TARGET RECOGNITION

A) Introduction

Different multiresolution hierarchical trees have been designed from different beginning decomposition levels, from 1 to 4. An example of tree built from the decomposition level 4, using Haar wavelet and K -means hard partitioning algorithm, is shown in Fig. 5.

This tree has 21 final clusters, an average distortion of 0.56, and a partition entropy of 2.9. In this figure, the clusters are designated by $C_{l,j}$, where l is the cluster number at resolution j . The number in each circle is the percentage of data in the cluster.

Total distortion and entropy are presented in Figs 6 and 7. A decrease of total distortion and an increase of entropy as a function of the number of clusters are observed which confirms the validity of the tree designing method.

In Figs 8 and 9, examples of range profiles contained in some clusters are shown. We can see that range profiles are associated according to aspect.

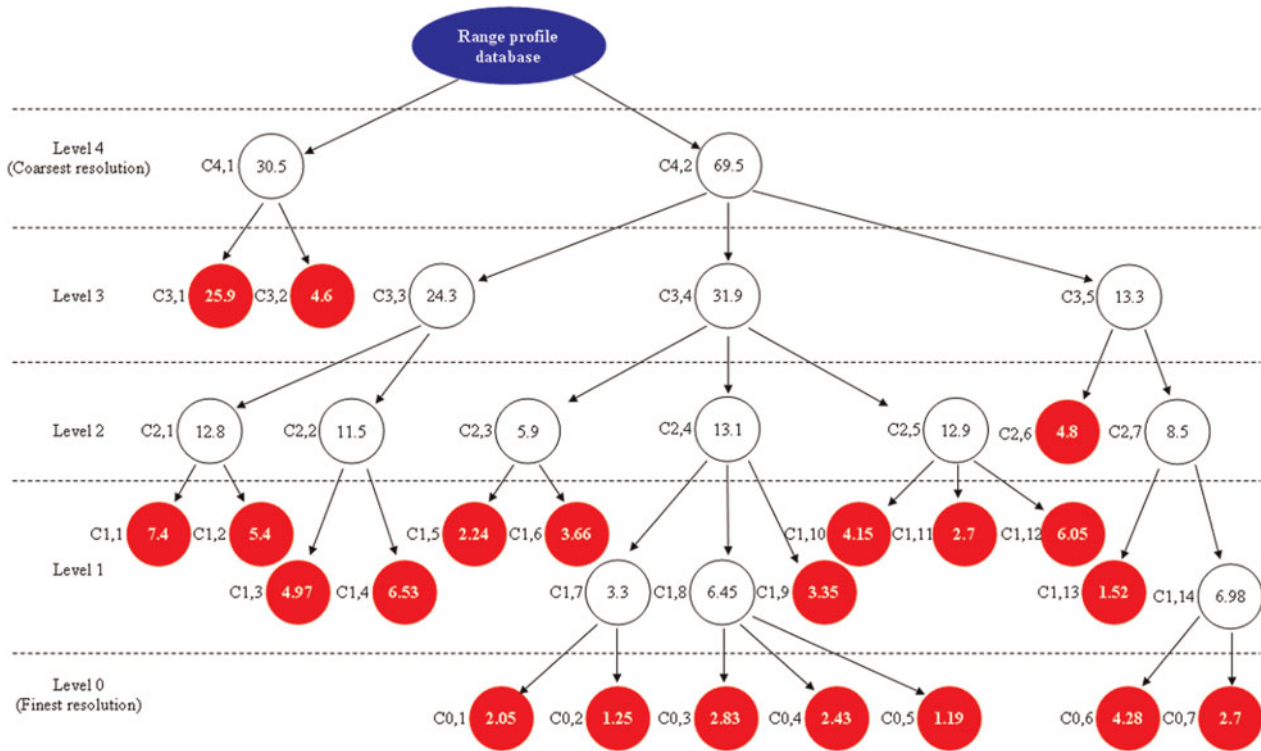


Fig. 5. Example of multiresolution hierarchical tree built from the level decomposition 4, using Haar wavelet and *K*-means clustering algorithm.

Figure 10 presents an estimation of *Pfc* as a function of *SNR* for different multiresolution hierarchical trees designed from different beginning decomposition levels, from 1 to 4. This figure reveals that higher is the beginning decomposition level, lower is the *Pfc*.

Figures 11 and 12 show, respectively, the variation of minimum *SNR* to obtain a *Pfc* smaller than 1%, and the search computational time *Sct* for a fixed *SNR* as a function of the beginning decomposition level used to design the multiresolution hierarchical tree. We observe a degradation of the minimum *SNR* to have a *Pfc* < 1% of 8 dB, but *Sct* is divided by a factor of 13.

Thus, multiresolution hierarchical trees are a good solution to compress high-resolution data of radar targets. Nevertheless, other methods exist such as the use of wavelet decomposition or

the unsupervised data clustering [15, 16]. It must be interesting to compare these techniques as a function of probability of false classification and computational time of search.

Finally, Figs 11 and 12 compare the efficiencies of different techniques (multiresolution hierarchical trees, *K*-means clustering algorithm, multiresolution Haar wavelet decomposition). The lowest *Sct* is obtained for clustering algorithm but with the most important degradation of minimum *SNR* to obtain a *Pfc* smaller than 1%. Use of approximation signals of wavelet decomposition to NCTR application makes it possible to obtain the weakest *SNR* to have a *Pfc* smaller than 1%, in particular for the first (finer) decomposition levels (1 and 2). Use of multiresolution hierarchical trees, designed from the coarser decomposition levels (3 and 4), is a good compromise between the data clustering and

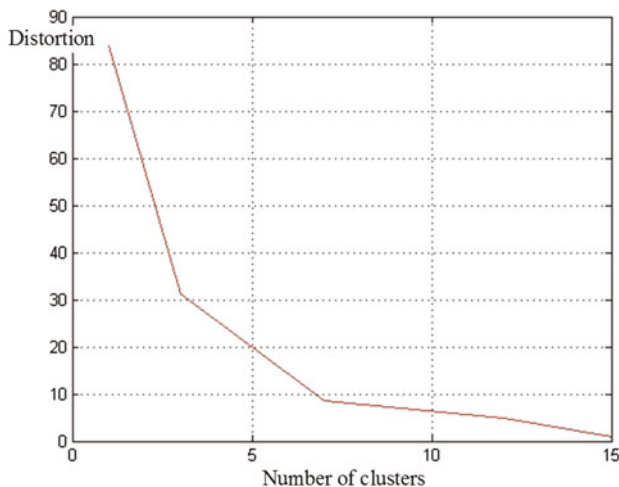


Fig. 6. Estimation of total distortion as a function of the number of clusters.

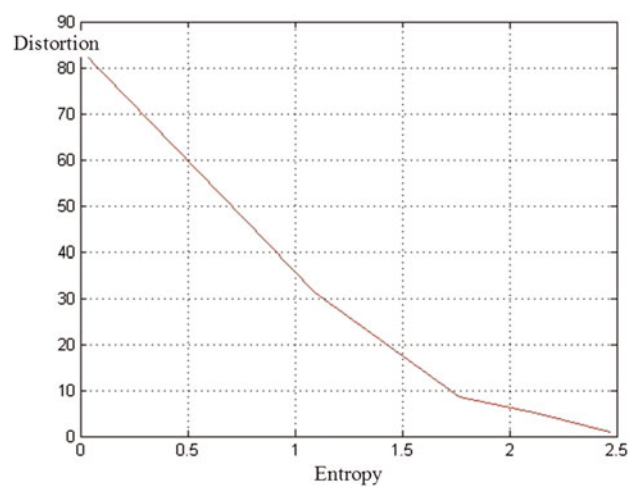


Fig. 7. Estimation of total distortion as a function of entropy.

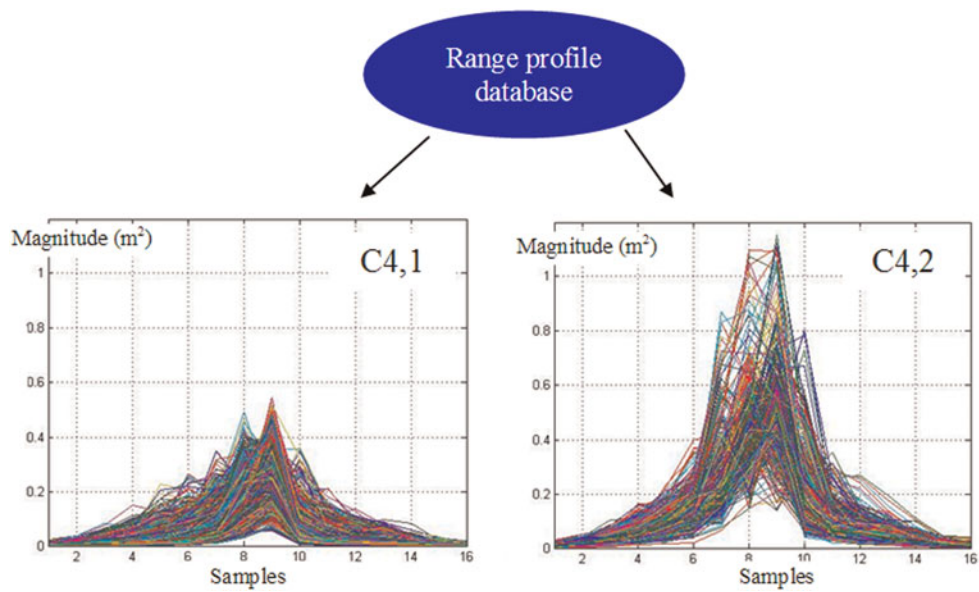


Fig. 8. Example of range profiles in each cluster at decomposition level 4.

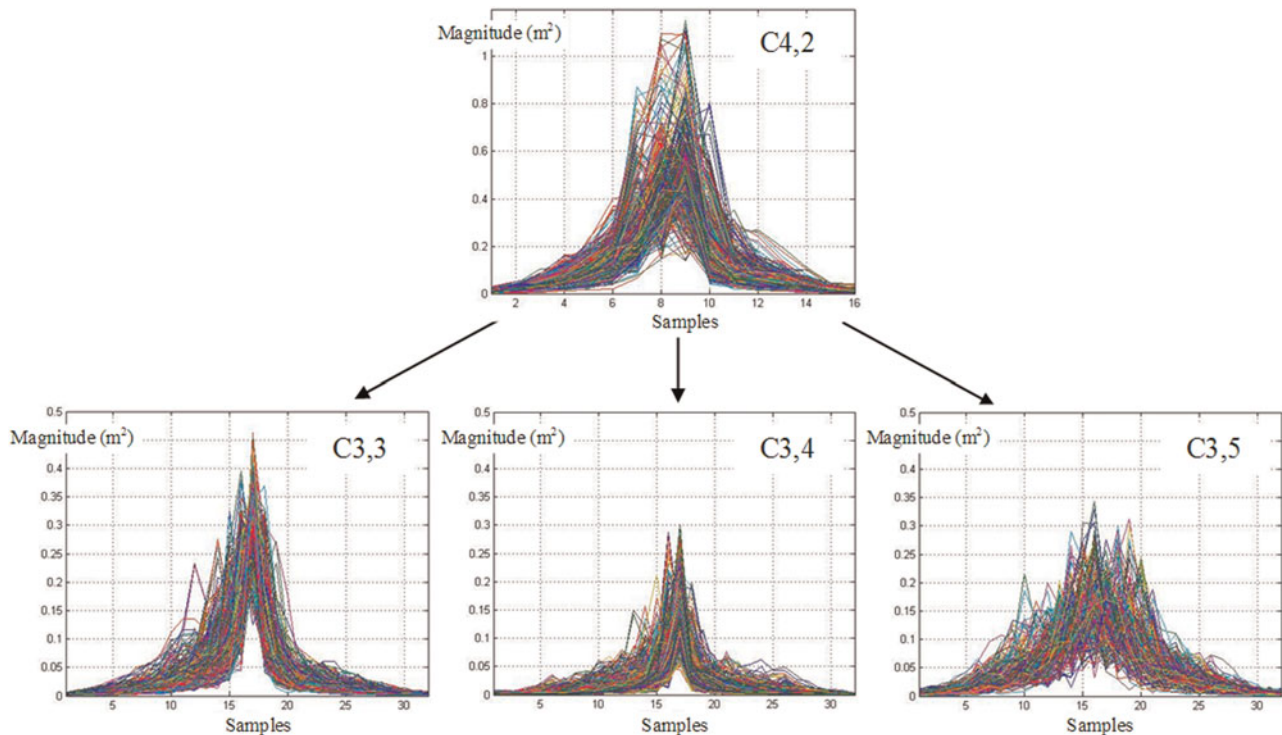


Fig. 9. Example of range profiles in each cluster at decomposition level 3.

the wavelet decomposition, because better performances are obtained for the minimum SNR to obtain a Pfc smaller than 1%, with a similar search computational time.

VII. CONCLUSION

The aim of this paper was to present a method to design tree structured hierarchical wavelet representations and to evaluate the efficiency of these trees to minimize the computational search time to NCTR association, compared to unsupervised

algorithms and multiresolution wavelet decomposition. The hierarchical designing method based on the use of approximation signals of the wavelet decomposition coupled with the K -means unsupervised clustering algorithm, is described. A criterion is presented to determine the number of clusters on each level of the tree with a hierarchical dependence.

For a hierarchical tree designed from the decomposition level 4, Sct is divided by a factor of 13, with a degradation of the minimum SNR to have a $Pfc < 1\%$ of 8 dB. Comparison with other database compression methods (wavelet decomposition, hard clustering) shows that multiresolution

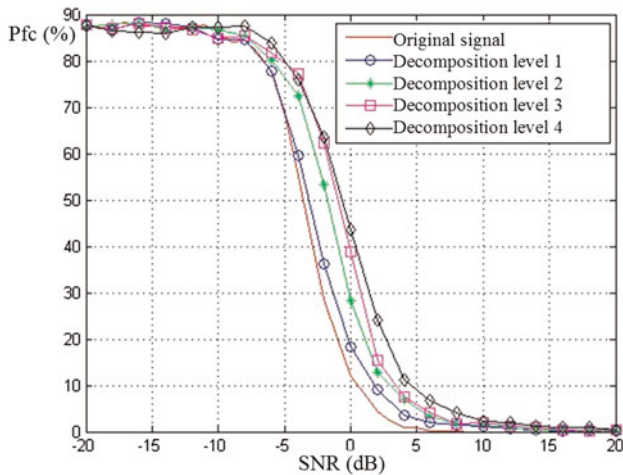


Fig. 10. Example of probability of false classification P_{fc} as a function of SNR for the original set and the multiresolution hierarchical trees designed from different beginning decomposition levels, from 1 to 4, using a Haar wavelet and K -means algorithm.

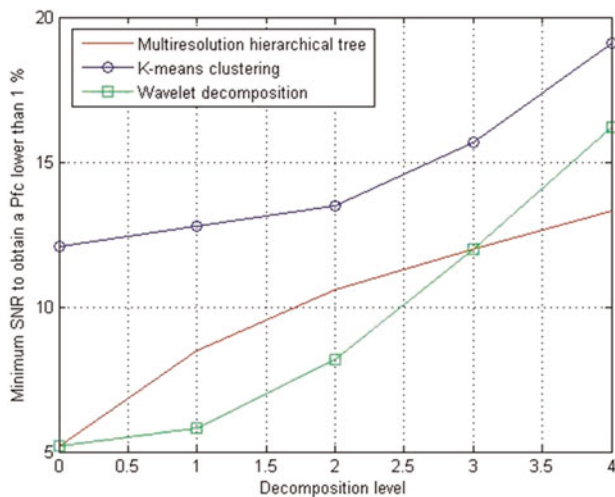


Fig. 11. Minimum SNR to obtain a P_{fc} smaller than 1% as a function of decomposition level for the multiresolution hierarchical trees, K -means algorithm ($C = 50$), and Haar wavelet decomposition.

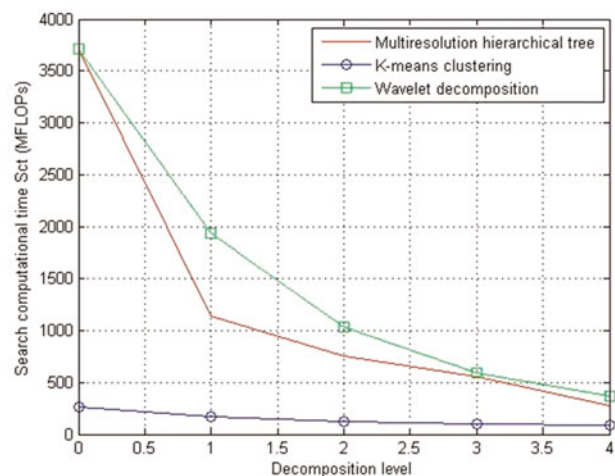


Fig. 12. Search computational time S_{ct} as a function of the beginning decomposition level for the multiresolution hierarchical trees, K -means algorithm ($C = 50$), and Haar wavelet decomposition.

hierarchical trees are a good compromise as a function of S_{ct} and P_{fc} , if they are designed from the upper (coarser) decomposition levels.

ACKNOWLEDGEMENTS

The author thanks THALES Air Systems and the French Ministry of Defence – DGA for their support to this study.

REFERENCES

- [1] Baras, J.S.; Dey, S.: Adaptive classification based on compressed data using learning vector quantization, in Proc. 38th Conf. on Decision & Control, Phoenix, AZ, USA, 1999.
- [2] Barès, C.; Brousseau, C.; Bourdillon, A.: A multifrequency HF–VHF radar system for aircraft identification, in IEEE Int. Radar Conf., Arlington, VA, USA, 2005.
- [3] Brousseau, C.: Définition, réalisation et tests d’un radar VHF multi-fréquence et multipolarisation – project MOSAR. PhD Thesis, University of Rennes I, 1995.
- [4] Burke, G.J.; Poggio, A.J.: Numerical electromagnetic code – method of moments, Part I: program description and theory. Technical Document 116, Naval Electronics Systems Command (ELEX 3041), 1977.
- [5] David, A.; Brousseau, C.; Bourdillon, A.: Simulations and measurements of radar cross section of a Boeing 747–200 in the 20–60 MHz frequency band. Radio Sci., **38**(4) (2003), 1064–1070.
- [6] Mallat, S.: A Wavelet Tour of Signal Processing, Academic Press, Paris, 1998.
- [7] Mallat, S.: A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell., **11** (7) (1989), 674–693.
- [8] MacQueen, J.B.: Some methods for classification and analysis of multivariate observations, in Proc. 5th Berkeley Symp. on Mathematical Statistics and Probability, Berkeley, CA, USA, 1967.
- [9] Zeng, Y.; Starzyk, J.: Statistical approach to clustering in pattern recognition, in Proc. 33rd Southeastern Symp. on System Theory, Athens, OH, USA, 2001.
- [10] Duda, R.O.; Hart, P.E.; Stork, D.G.: Pattern Classification, John Wiley & Sons, 2nd ed., New York, 2001.
- [11] Marques de Sà, J.P.: Pattern Recognition – Concepts, Methods and Applications, Springer Editions, 2001.
- [12] Bezdek, J.C.; Dunn, J.C.: Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. IEEE Trans. on Computers, **24** (1975), 835–838.
- [13] Dunn, J.C.: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, J. Cybern., **3**(3) (1973), 32–57.
- [14] Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [15] Brousseau, C.: Application of the multiresolution wavelet representation to non-cooperative target recognition, in Radar’09 – Int. Radar Conf., Bordeaux, France, 2009.
- [16] Brousseau, C.: Estimation of efficiency of multiresolution approach applied to non-cooperative target recognition, in Radar2010 – IEEE Int. Radar Conf., USA, 2010.
- [17] Cover, T.; Hart, P.: Nearest neighbour pattern classification. IEEE Trans. Inf. Theory, **IT-13** (1) (1967), 21–27.



Christian Brousseau was born in France in 1966. He received his Ph.D. degree in telecommunications and signal processing from the University of Rennes 1 in 1995. His main research interests are development of new concepts of radar system applied to non-cooperative target recognition, and study and modeling propagation

phenomenon and their effects on radio communication systems and radar systems.