

KNOWLEDGE-BASED RECURSIVE LEAST SQUARES TECHNIQUES FOR HETEROGENEOUS CLUTTER SUPPRESSION

Antonio De Maio, Alfonso Farina, and Goffredo Foglia

A. De Maio is with Università degli Studi di Napoli "Federico II", Via Claudio 21, I-80125 Napoli, Italy; e-mail: a.demaio@unina.it

A. Farina is with Selex - Sistemi Integrati Via Tiburtina Km.12.4, 00131, Roma, Italy; e-mail: afarina@selex-si.com.

G. Foglia is with Elettronica S.p.A., Via Tiburtina Km.13.7, 00131, Roma, Italy; e-mail: goffredo.foglia@elt.it.

ABSTRACT

In this paper we deal with the design of Knowledge-Based adaptive algorithms for the cancellation of heterogeneous clutter. To this end we revisit the application of the Recursive Least Squares (RLS) technique for the rejection of unwanted clutter and devise modified RLS filtering procedure accounting for the spatial variation of the clutter power. Then we introduce the concept of Knowledge-Based RLS and explain how the a-priori knowledge about the radar operating environment can be adopted for improving the system performance. Finally we assess the benefits resulting from the use Knowledge-Based processing both on simulated and on measured clutter data collected by the McMaster IPIX radar in November 1993.

1. INTRODUCTION

Adaptive filters for clutter suppression are sub-optimal implementations of the optimum linear processor which, assuming the a-priori knowledge of the disturbance spectral properties, maximizes the output Signal to Interference plus Noise power Ratio (SINR) [1]. In order to estimate the clutter covariance matrix adaptive filters exploit training data, namely clutter returns collected from range cells spatially close to the one under test and assumed free of useful signal component. The estimation procedure is usually performed through the sample covariance matrix and then the adaptive filter weights are derived [2]. It is clear that the imperfect estimate of the clutter Power Spectral Density (PSD) leads to sub-optimal performances even in the presence of a spatially homogeneous environment, namely secondary data independent and identically distributed (iid).

Nevertheless the assumption of homogeneous clutter over the extent of the reference window is restrictive and quite often violated. Clutter heterogeneities are usually present yielding severe losses in the performance of adaptive filters [3].

A possible approach to circumvent the severe performance loss caused by clutter heterogeneities is to resort to Knowledge-Based techniques, which should be valuable in using a-priori information to restore the radar performance [4]. Examples of a-priori knowledge are Digital Terrain Elevation Models (DTEM), previous look data, Geographic Information Systems (GIS's), roadways (to highlight sectors of surveillance where moving cars or vehicles might be present), background of air/surface traffic, system calibration

information, et cetera. The ultimate goal is to make the radar an intelligent device, such that it is capable of developing cognition of the surrounding environment [5].

In this paper we show how Knowledge-Based techniques can be exploited for the adaptive implementation of the optimum doppler processor. To this end, we first revisit the Recursive Least Squares (RLS) algorithm [6], which permits a reduced complexity adaptive implementation of the optimum filter, designed under the assumption of homogeneous environment. Then we introduce a recursive procedures exploiting data normalizations. However, due to the strong clutter variability, none of the algorithms uniformly outperforms the others and the problem of choosing the most appropriate procedure arises. In this context we propose to use the environmental information for the selection of the most suitable adaptation technique and to mitigate the deleterious effects of clutter heterogeneity.

2. PROBLEM FORMULATION AND DESIGN ISSUES

Denote by r the N -dimensional complex vector of the samples from the base-band equivalent of the received signal from the cell under test. Under hypothesis H_0 , namely target absence, r contains disturbance only, i.e.

$$H_0 : \quad r = d,$$

where the disturbance vector d accounts for both clutter and thermal noise. Under H_1 , instead, r also contains a target component, i.e.

$$H_1 : \quad r = \alpha p + d,$$

with α the complex amplitude accounting for both the target as well as the channel propagation effects and p the target temporal steering vector.

A doppler processor performs the inner product between a suitable weight vector w and the vector r of the returns from the cell under test. Ideally it provides coherent gain on target while forming doppler response nulls to suppress the disturbance components. Specifically the optimum doppler filter, which maximizes the output SINR [7], is given to within a scale factor by $w = M^{-1} p$, where M denotes the disturbance covariance matrix: the design of the optimum filter requires a-priori knowledge of the clutter covariance matrix. From a practical point of view, this knowledge is

not available, and thus it is necessary to develop a processor which online estimates M and automatically adapts its filtering action in response to a changing environment.

In order to perform the covariance estimate, it is customary to resort to training data, namely returns from range cells spatially close to the one under test and free of useful signal components. Moreover, it is assumed that they share the same covariance matrix of the disturbance component from the cell under test (homogeneous assumption). Denoting by r_1, \dots, r_K the N -dimensional vectors of the training set ($K \geq N$) the Maximum Likelihood (ML) estimator of M is the sample covariance matrix

$$\hat{M} = \frac{1}{K} \sum_{k=1}^K r_k r_k^\dagger, \quad (1)$$

where $(\cdot)^\dagger$ denotes conjugate transpose.

The adaptive filter which use the sample covariance matrix (1) require the computation of \hat{M} and its online inversion. It is well known that the computational burden connected with this last operation is $O(N^3)$. Hence, in order to save computation time, recursive algorithms which estimate the inverse covariance directly from the input samples can be conceived.

2.1 RLS Covariance Update

The RLS algorithm was first conceived for beamforming applications in [8] and then adopted for clutter suppression purposes in [9]. It relies on a recursive procedure which estimates the inverse covariance matrix M^{-1} directly from the input samples. By doing so the online matrix inversion is no longer required, and a significant saving in computation time can be achieved.

The recursive equation which defines the RLS can be obtained by applying the matrix inversion lemma [10] to the equation

$$\hat{M}_{k+1} = \lambda \hat{M}_k + r_{k+1} r_{k+1}^\dagger, \quad 0 < \lambda \leq 1, \quad (2)$$

where the weighting coefficient λ , referred to in the sequel as the forgetting factor, determines the relative confidence of the input data with respect to the current estimate. The aforementioned procedure leads to

$$\hat{M}_{k+1}^{-1} = \lambda^{-1} \hat{M}_k^{-1} - \lambda^{-2} \frac{\hat{M}_k^{-1} r_{k+1} r_{k+1}^\dagger \hat{M}_k^{-1}}{1 + \lambda^{-1} r_{k+1}^\dagger \hat{M}_k^{-1} r_{k+1}}, \quad (3)$$

which defines the estimate of the inverse covariance at the step $k+1$ as a function of the estimate at the previous step, the new input data r_{k+1} vector, and the forgetting factor.

This last parameter rules the speed of adaptation of the algorithm. In fact, the higher λ the more accurate the estimate. This, however, implies a longer reaction time due to the longer memory of the algorithm. A faster adaptation is achieved by means of a smaller forgetting factor at the expense of a poorer accuracy, resulting in a loss of the steady state SINR. As a consequence, a trade off between speed and accuracy must be achieved.

The recursion is usually initialized assuming $\hat{M}_0 = \delta I$ (I is the identity matrix and δ is a scalar called loading factor) which is tantamount to introducing in the covariance estimate

an exponentially decaying diagonal loading. The computational complexity connected with the implementation of the RLS is $O(N^2)$ flops, which indicates a significant saving in computation time with respect to the online inversion of the sample covariance matrix.

2.2 Normalized RLS (NRLS) Covariance Update

In the presence of a heterogeneous environment, where the clutter power varies from cell to cell, the previous filter suffers a severe performance loss with respect to the optimum processor. This behavior is due to the fact that in this situation the sample covariance matrix is no longer the ML covariance estimator. A possible way to restore the system performance in such a case is to employ a normalized sample covariance matrix [11] in place of the usual one, i.e.

$$\hat{M}_n = \frac{N}{K} \sum_{k=1}^K \frac{r_k r_k^\dagger}{\|r_k\|^2}, \quad (4)$$

where $\|\cdot\|$ denotes the Euclidean norm of a complex vector. A normalized version of the RLS algorithm can also be conceived. Precisely, denoting by $r_{k,n} = \frac{r_k}{\|r_k\|/\sqrt{N}}$, the k -th normalized data vector, we get the following equation for the recursive update of the inverse covariance matrix

$$\hat{M}_{k+1,n}^{-1} = \lambda^{-1} \hat{M}_{k,n}^{-1} - \lambda^{-2} \frac{\hat{M}_{k,n}^{-1} r_{k+1,n} r_{k+1,n}^\dagger \hat{M}_{k,n}^{-1}}{1 + \lambda^{-1} r_{k+1,n}^\dagger \hat{M}_{k,n}^{-1} r_{k+1,n}}, \quad (5)$$

where $\hat{M}_{k,n}^{-1}$ denotes the estimate at the step k .

The computational complexity connected with the implementation of the NRLS is $O(N^2)$ flops.

2.3 Knowledge-Based RLS (KB-RLS) Covariance Update

The previous subsections highlight that two different updating algorithms for the filter coefficients can be conceived. The former exploits a conventional recursive covariance estimator whereas the latter resorts to a normalized one. It is clear that in the presence of a homogeneous environment the RLS provides better performance than the NRLS, since in this situation the conventional sample covariance matrix is the ML estimate of the clutter covariance. On the contrary, if the scene is such that the clutter power varies from one cell to another, then the NRLS usually outperforms the classic RLS procedure. It follows that the problem of selecting the most suitable filtering procedure arises. To this end it is quite natural to exploit Knowledge-Based techniques, which might be very valuable in using a-priori information to select the most appropriate filtering algorithm. In this context, exploiting the geographical information provided by a GIS, it is possible to know the exact location of transition regions where the updating algorithm must be changed. Otherwise stated, if we refer to an interface between homogeneous land (region 1) and sea (region 2), resorting to the GIS information, we can use in region 1 the RLS procedure while in region 2 the NRLS algorithm.

Another relevant source of a-priori information which can be exploited for the selection of the algorithm is the wind data. Precisely, if the filter must operate in a homogeneous region covered by vegetation, it is reasonable to employ the

plain RLS algorithm in the absence or for weak wind. However, for very windy conditions (for instance in the presence of wind gusts), it is known that the windblown vegetation causes strong fluctuations of the clutter power from cell to cell [12] and thus the NRLS must be adopted.

We also point out that Knowledge-Based techniques might be employed for the choice of the forgetting factor λ . In other words, we suggest to adopt a space-varying forgetting factor whose value at the step k must be a function of the environment, i.e. $\lambda = \lambda_k$ (environment).

The aforementioned idea may represent a very powerful technique to mitigate the effects of clutter edges. Indeed, exploiting a GIS for locating the presence of transition regions, we can think to lower (according to a certain mathematical law) the forgetting factor when we are approaching the first side of the interface, and then to gradually increase λ right after the transition. By doing so, as it will be shown in the performance analysis, a-priori knowledge helps the algorithm to rapidly forget the clutter conditions within the first region. Otherwise stated, the filter will exhibit a shorter reaction time, namely a faster adaptation, than the plain RLS.

3. PERFORMANCE ANALYSIS

This section is devoted to the performance assessment of the previously presented algorithms both on simulated data and on real radar data (X-band mixed sea and land clutter) which was collected by the McMaster University IPIX radar in November 1993.

All the experiments assume that

- the number of integrated pulses N is equal to 16;
- the structure of the useful signal component is

$$p = \frac{1}{\sqrt{N}} [1, \exp(j2\pi f_d), \dots, \exp(j2\pi(N-1)f_d)],$$

where $f_d = 0.01$ (slow moving target) is the normalized doppler frequency;

- the loading factor δ for the initialization of the algorithms is equal to 10^3 which is tantamount to assume $w_0 = \frac{1}{\delta} p$.

3.1 Simulated Data

In the first experiment (Figure 1a) we assume a homogeneous Gaussian environment. Namely, we model the vectors r_k , $k = 1, \dots, K$, as zero-mean iid complex circular Gaussian vectors with covariance matrix $M = M_c + \sigma^2 I$ where M_c is the clutter covariance and σ^2 is the thermal noise level. The matrix M_c is assumed exponentially shaped, namely its (i, j) -th entry is given by

$$M_c(i, j) = P_c \rho^{|i-j|} \exp(j2\pi(i-j)f_c) \quad (6)$$

where ρ is the one-lag correlation coefficient, and P_c and f_c denote respectively the clutter power level and doppler frequency. We also suppose $\rho = 0.999$, $f_c = 0$, and $\text{CNR} = \frac{P_c}{\sigma^2} = 30$ dB which is the typical situation of a homogeneous land environment. The performance is evaluated in terms of SINR, i.e.

$$\text{SINR} = \frac{|w_k^\dagger p|^2}{w_k^\dagger M w_k}, \quad (7)$$

at the k -th instance of the recursion. Precisely, in Figure 1a,

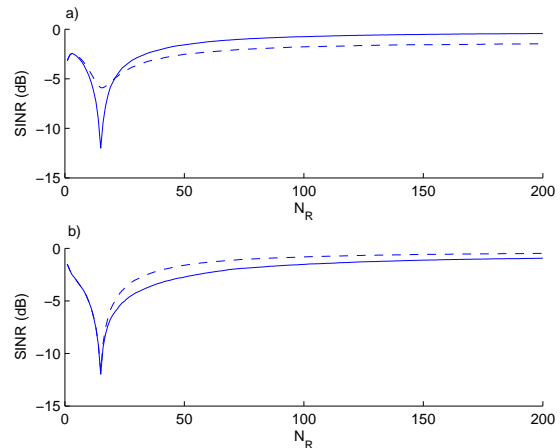


Figure 1: Normalized SINR (dB) versus the number of range samples N_R for the RLS (solid line) and the NRLS (dashed line). a) Simulated land clutter. b) Simulated sea clutter.

the normalized SINR of the RLS and NRLS (both with forgetting factor $\lambda = 0.99$), averaged over 400 independent trials, is plotted versus the number of range samples N_R . As expected, since the simulated environment is homogeneous, after a transient of about 30 cells, the plain RLS outperforms the normalized algorithm. An opposite behavior is shown by the plots of Figure 1b where a heterogeneous scenario accounting for the spatial variations of the clutter power is considered. Therein the disturbance vectors r_k are modeled as the sum of a clutter contribution c_k and the receiver thermal noise n_k (statistically independent of c_k), i.e. $r_k = c_k + n_k$. The n_k are iid zero-mean complex circular Gaussian vectors with covariance matrix $\sigma^2 I$. The clutter vectors c_k are iid complex circular Spherically Invariant Random Vectors (SIRV's) [13] with covariance matrix given by (6) and Gamma distributed squared textures whose shape parameter, which rules the impulsive nature of the clutter, is denoted by ν . The plots also assume $\rho = 0.8$, $f_c = 0.2$, $\nu = 0.5$, and $\text{CNR} = 20$ dB, which is a typical scenario of sea clutter. Finally, the forgetting factor is again $\lambda = 0.99$. It is clear that the NRLS outperforms the plain RLS and this behavior stems from the observation that in the simulated environment the sample covariance matrix is no longer the ML estimator of the disturbance covariance.

The previous analyses have shown that no algorithm uniformly outperforms the others and the superiority of a specific filtering procedure is strictly dependent on the actual disturbance environment. Thus it appears quite natural to exploit a-priori knowledge, provided by a GIS about the observed geographical site, in order to select the most appropriate clutter suppression algorithm.

In the next example the performance improvements achievable resorting to Knowledge-Based processing are illustrated. Therein a heterogeneous scenario composed of two different regions is simulated. The former, composed of the range cells from $k = 1$ to 400, contains homogeneous land clutter plus noise (simulated according to the model used in Figure 1a with $\rho = 0.999$, $f_c = 0$, and $\text{CNR} = 30$ dB). The latter, namely range cells from $k = 400$ to 800, includes sea clutter plus thermal noise (simulated according to the model used in Figure 1b with $\rho = 0.8$, $f_c = 0.2$, and $\text{CNR} = 20$ dB).

The RLS and NRLS with forgetting factor $\lambda = 0.99$ are compared with the KB-RLS that exploits a-priori information concerning the location of the clutter edge. Precisely, the

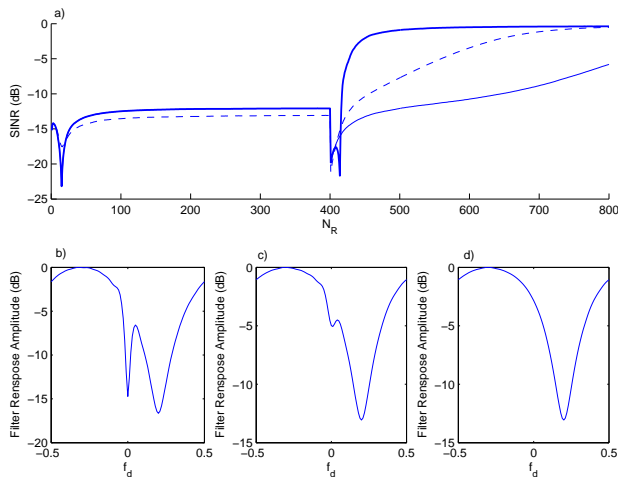


Figure 2: a) Normalized SINR (dB) versus the number of range samples N_R for the RLS (solid line), the NRLS (dashed line), and the KB-RLS (bold line). Simulated land clutter for $1 \leq N_R \leq 400$. Simulated sea clutter for $N_R > 400$. Amplitude of the filter frequency response (dB) versus the normalized doppler frequency f_d for the RLS (b), the NRLS (c), and the KB-RLS (d).

KB-RLS coincides with the plain RLS algorithm in the first region. Moreover, after the clutter edge, it switches into the NRLS. As to the forgetting factor, the KB-NRLS algorithm assumes a space-varying λ in order to provide a fast adaptation to the new clutter conditions. Otherwise stated, the forgetting factor is ruled by the following linear piecewise function $\lambda_k(\text{environment}) = 0.99, \forall k \in [1, 400] \cup [421, 800]$ and $\lambda_k(\text{environment}) = 0.5 + 0.0245(k - 401), \forall k \in [401, 420]$. When, in correspondence of the clutter edge, the KB-RLS commutes into the NRLS the inverse covariance estimate is normalized, i.e. $\hat{M}_{401}^{-1} = (1 - \lambda)\text{tr}(\hat{M}_{400})\hat{M}_{400}^{-1}$, with $\text{tr}(\cdot)$ denoting the trace of a square matrix. Extensive simulation results have shown that this normalization ensures a fast convergence even if it adds a small increase in the algorithm computational complexity, since a further scalar recursive equation for the trace updating, i.e. is required.

In Figure 2a, the normalized averaged SINR of the RLS, NRLS, and KB-RLS is plotted versus N_R . The curves highlight an abrupt transition in correspondence of the clutter edge. The RLS performs better than NRLS in the first region while the opposite behavior is observed in the second region. The KB-RLS algorithm achieves the best performance, especially after the transition, requiring a shorter adaptation time than the NRLS to reach the steady state SINR. This behavior is confirmed by the curves of Figure 2b where the amplitude of the filter frequency response at range cell 800 is plotted versus the normalized frequency. The plots show that the RLS does not yet forget the land clutter environment; indeed the filter response exhibits a deep null in correspondence of $f_c = 0$. The quoted null is less pronounced with reference to the curve of the NRLS and vanishes if the KB-NRLS is employed. In this last case, only one null is present in correspondence of $f_c = 0.2$, namely the clutter doppler frequency in the second region.

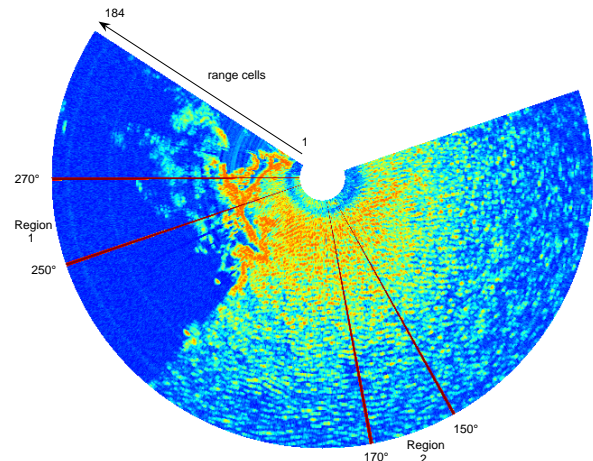


Figure 3: 2-D intensity field of the mixed land and sea clutter live data (red are the strongest returns, blue are the weaker returns). The straight red lines delimitate the selected regions, namely Region 1 ($[250^\circ, 270^\circ]$, land clutter) and Region 2 ($[150^\circ, 170^\circ]$, sea clutter).

3.2 Real Data

Radar measurements were collected in November 1993 using the McMaster IPIX radar from a site in Dartmouth [14], Nova Scotia, on the East Coast of Canada. The radar was mounted on a cliff facing the Atlantic Ocean, at a height of 100 feet above the mean sea level, and scans the site over 370 deg in 10 seconds in a continuous azimuth scan mode. More details on the experiment (Dataset13) can be found in [14]. The 2-D clutter intensity field is plotted in Figure 3 showing the presence of two different sectors containing respectively returns from land and sea. In order to perform the analysis in the presence of clutter edge, we select two different angular sectors and put them near. Precisely, the selected areas are

- Region 1, from 250 to 270 deg, where there is a land clutter region including the lands near Halifax.
- Region 2, from 150 to 170 deg, where there is an open view of the Atlantic Ocean.

We employ $N = 16$ azimuth returns and a total of 400 trials for estimating the SINR in a given range position. The RLS and NRLS exploit a forgetting factor equal to 0.99. The KB-RLS coincides with the RLS algorithm but exploits a space-varying forgetting factor ruled by the linear piecewise function $\lambda_k(\text{environment}) = 0.99, \forall k \in [1, 184] \cup [295, 386]$ and $\lambda_k(\text{environment}) = 0.5 + 0.0245(k - 185), \forall k \in [185, 204]$. In Figure 4 the normalized SINR is plotted versus N_R showing that the NRLS provide a faster convergence rate than the RLS in the first 50 cells but is outperformed by the KB-RLS, until the end of the Region 1. After the clutter edge the NRLS perform better than the non-normalized algorithm. As to the KB-RLS, it requires a short adaptation time after the clutter edge, achieving its best performance from cell 225 to the end of Region 2. A further evidence of this behavior is highlighted in Figure 5, where the amplitude of the filter frequency response at the end of Region 2 is plotted versus the normalized doppler frequency. Figure 5a refers to the RLS filter and shows a null around zero doppler frequency. It is due to the filter memory which has not yet forgotten the clutter condition of the first region and its effect is the degradation of the performance achievable for slow moving targets.

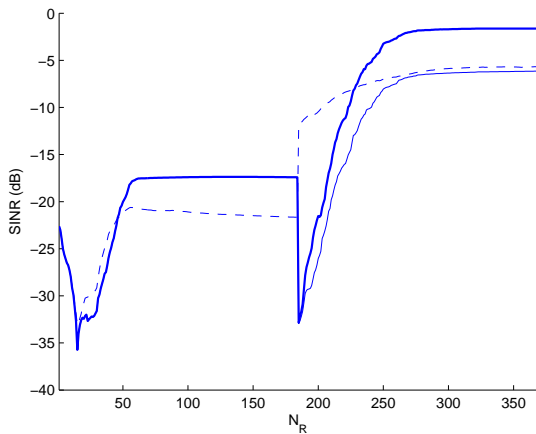


Figure 4: Normalized SINR (dB) versus the number of range samples N_R for the RLS (solid line), the NRLS (dashed line), and the KB-RLS (bold line). Real land clutter (Region 1) for $1 \leq N_R \leq 184$. Real sea clutter (Region 2) for $N_R > 184$.

On the contrary, the shape of the KB-RLS filter (displayed in Figure 5b) is correctly adapted to the sea clutter PSD.

4. CONCLUSIONS

In this paper we have considered the design and the analysis of RLS algorithms for the suppression of heterogeneous clutter. First of all we have revisited the standard RLS procedure and then we have introduced a normalized RLS algorithm. Since there exists no technique which uniformly outperforms the other, we have investigated the use of a-priori information for the choice of the most suitable clutter suppression procedure. This idea leads to the concept of KB-RLS where the environmental conditions, in particular GIS data, are extensively exploited for the improvement of the system performance.

We have conducted a performance analysis both in the presence of simulated and measured radar data, collected by the McMaster IPIX radar in 1993, showing the effectiveness of the Knowledge-Based approach. Precisely, if a-priori information is suitably exploited, then significant performance improvements can be achieved.

Possible future research tracks include the analysis of the KB-RLS in the presence of other real datasets, collected by both ground-based and airborne radars, as well as the problem of devising the optimum function for the forgetting factor variations.

ACKNOWLEDGMENT

The authors are deeply indebted to Prof. S. Haykin and Dr. B. Currie of the McMaster University (Canada) who have kindly provided the IPIX data.

REFERENCES

- [1] R. A. Monzingo and T. W. Miller, "Introduction to Adaptive Arrays," John Wiley & Sons, 1980.
- [2] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 10, No. 4, pp. 853-863, November 1974.
- [3] R. Nitzberg, "An Effect of Range Heterogeneous Clutter on Adaptive Doppler Filters," *IEEE Trans. on Aerospace*

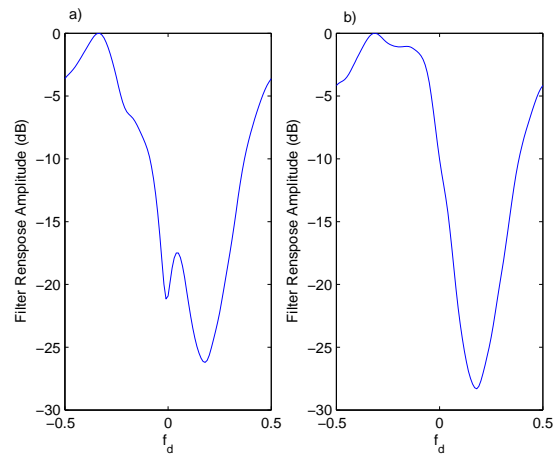


Figure 5: Amplitude of the filter frequency response (dB) versus normalized doppler frequency f_d for the RLS (a) and the KB-RLS (b).

and Electronic Systems, Vol. 26, No. 3, pp. 475-480, July 1990.

- [4] W. L. Melvin, M. Wicks, P. Antonik, Y. Salama, P. Li, and H. Schuman, "Knowledge-Based Space-Time Adaptive Processing for Airborne Early Warning Radar," *IEEE AES Systems Magazine*, Vol. 13, No. 4, pp. 37-42, April 1998.
- [5] S. Haykin, "Radar Vision," *Proc. of the Second International Specialist Seminar on the Design and Application of Parallel Digital Processors*, pp. 75-78, April 1991.
- [6] J. E. Hudson, "Adaptive Array Principles," *Peter Peregrinus*, New York and London, 1981.
- [7] L. E. Brennan and I. S. Reed, "Theory of Adaptive Radar," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 9, No. 1, pp. 237-252, March 1973.
- [8] E. B. Lunde, "The Forgotten Algorithm in Adaptive Beamforming," published in "Aspects of Signal Processing," G. Tacconi (Ed), *Proc. of the NATO Advanced Study Institute on Signal Processing and Underwater Acoustic*, Porto Venere, Italy, 1976.
- [9] A. Farina, F. A. Studer, and E. Turco, "Adaptive Methods to Implement the Optimum Radar Signal Processor," *Proc. of International Conference on Radar*, pp. 42-47, Bangalore, India, October 1983.
- [10] R. A. Horn and C. R. Johnson, "Matrix Analysis," *Cambridge University Press*, 1985.
- [11] E. Conte, M. Lops, and G. Ricci, "Adaptive Detection Schemes in Compound-Gaussian Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 34, No. 4, pp. 1058-1069, October 1998.
- [12] J. B. Billingsley, "Low-Angle Radar Land Clutter," *Scitech Publishing*, 2002.
- [13] E. Conte and M. Longo, "Characterisation of Radar Clutter as a Spherically Invariant Random Process," *IEE Proc. Pt. F*, Vol. 134, No. 2, pp. 191-197, April 1987.
- [14] A. Drosopoulos, "Description of the OHGR Database," *Tech. Report No. 94-14*, Defence Research Establishment Ottawa, December 1994, Ottawa.