Akademie věd české republiky
ÚSTAV TEORIE INFORMACE A AUTOMATIZACE

Academy of Sciences of the Czech Republic
INSTITUTE OF INFORMATION THEORY AND AUTOMATION

RESEARCH REPORT

J.Kadlec

Structure Determination and Tracking for Parallel Radial Basis Function based Nonlinear Networks

No. 1790 January 1994

ÚTIA AV ČR P. O. Box 18, 182 08 Prague, Czech Republic
Telex: 122018 atom c, Fax: (+42)(2)6641 4903
e-mail: utia@utia.cas.cz

This report constitutes an unrefereed manuscript which is intended to be submitted for publication. Any opinions and conclusions expressed in this report are those of the author(s) and do not necessarily represent the views of the Institute.

Structure Determination and Tracking for Parallel Radial Basis Function based Nonlinear Networks

J. Kadlec

Institute of Information Theory and Automation,
Czech Academy of Sciences, P.O. Box 18, 182 08 Prague 8,
Czech Republic, e-mail: adapt@cspgas11.bitnet

Abstract. The paper proposes the concept of an array for structural estimation of nonlinear systems modelled by Radial Basis Functions. The proposed systolic array incorporates estimation of different modes in parallel and provides an estimate of the probabilities of these models. The array can be interpreted as a neural network type architecture, based on the systolic square–root, recursive least–squares algorithm.

1. Introduction

Feedforward layered neural networks have increasingly been used in many areas of nonlinear system identification and filtering [1], [2]. They model the nonlinearities by using sigmoidal functions and the weights (Fig. 1.) are obtained using nonlinear optimization.

1.1. Radial basis function model

An alternative to sigmoidal functions and nonlinear optimization is to use Radial Basis Functions (RBF) with fixed centres and linear optimization to estimate the weights [3]. The nonlinear system

$$\hat{y}(k) = \sum_{i=1}^{m} \hat{w}_i(k)\varphi_i(k) \tag{1}$$

is modelled by summing m radial basis functions which have variable weights and fixed centres. In (1) k is the discrete time and $\hat{w}_i(k)\,;\, i = 1, 2, ..., m$ is a set of real coefficients or weights to be determined or recursively estimated from the input - output measurements.

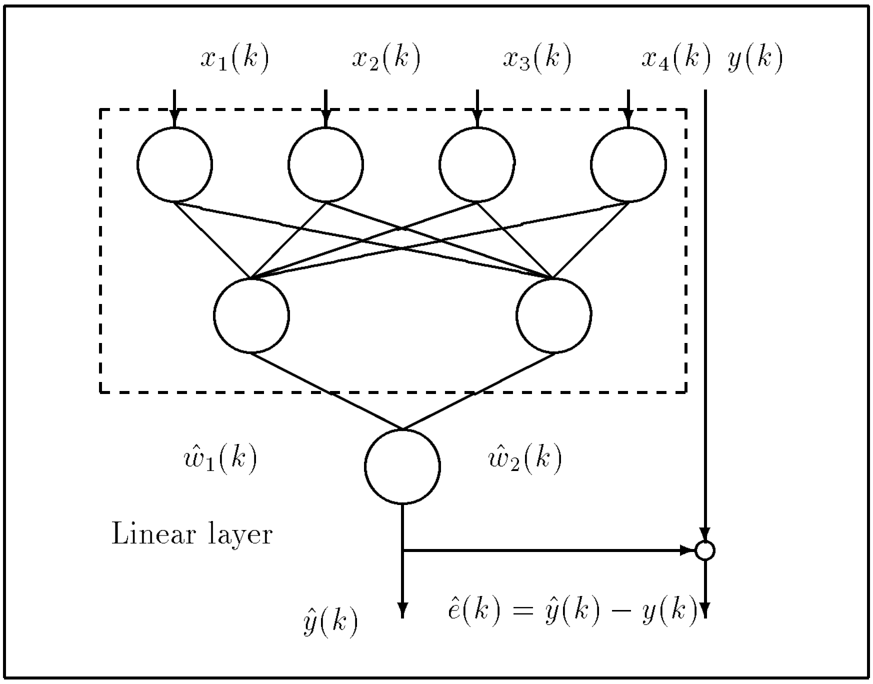

Figure 1. Feedforward layered neural network with the linear top layer.

In the RBF model the data $\varphi_i(k)$; $i = 1, 2, ..., m$ are constructed by a nonlinear transformation from the vector of measurements $X(k)$. The transformation is given by

$$\varphi_i(k) = \varphi \left\{ \parallel X(k) - X_i^c \parallel \right\} \; = \; \varphi \left\{ r_i(k) \right\} \tag{2}$$

where X_i^c ; $i = 1, 2, ..., m$ is a given set of "centre" vectors in the N_c – dimensional data space. The given nonlinear scalar functions

$$\varphi_i(k) = \varphi \left\{ r_i(k) \right\} \; ; \; i = 1, 2, ..., m \tag{3}$$

have as the argument the N_c–dimensional Euclidean distances $r_i(k) \geq 0$ between the data vector $X(k)$ and the i-th radial basis centre X_i^c. $\varphi_i(k)$ is therefore referred to as a radial basis function. The function

$$\varphi \left\{ r_i(k) \right\} = e^{-r_i(k) r_i(k)} \tag{4}$$

has been found [5] to give very good results, but many other function (exponential, cubic ...) can be used.

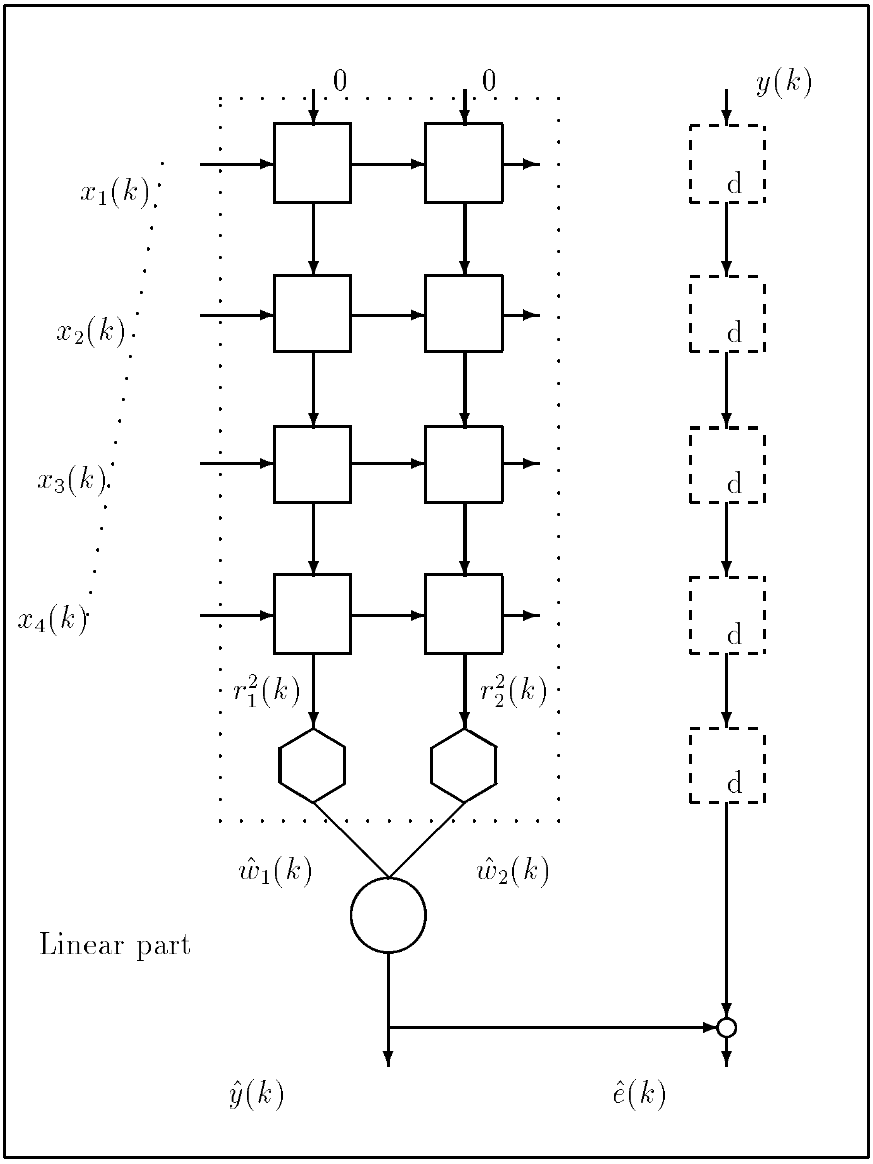

Figure 2. Parallel implementation of the radial basis function model.($m = 2$; $N_c = 4$)

Fig. 2. presents an array for computation of (1). The columns of processors compute $r_i^2(k) = \| X(k) - X_i^c \|^2$. Hexagonal blocks implement the RBF nonlinearities $\varphi\{r_i(k)\}$. The dashed blocks represent a delay. The slanted input reflects the pipelined computation.

2. Estimation as propagation of probability

In the time variable environment a system may have different modes of operation and therefore different RBF models. This is often indicated by multiple local maximum in the off-line searching for the values of optimal set of RBF's. We propose to handle these cases, by the concept of the Bayesian estimation.

2.1. Problem formulation and notation.

Our incomplete knowledge of the system behavior is formally expressed through a vector of m unknown, time-varying parameters $w(k) = [w_1(k), ..., w_m(k)]'$ characterised by the probability density function (p.d.f.) $p(w(k) \,|\, H_j,\, k-1)$. The abbreviation $p(. \,|\, H_j,\, k-1)$ denotes the conditioning on the input–output data measurement up to the time $k-1$ and the hypothesis H_j about the position of the radial basis centre.

The complete recursion of the adaptive estimator can be written in two steps. The first step is the data update $p(w(k) \,|\, H_j,\, k-1) \rightarrow p(w(k) \,|\, H_j,\, k)$. It forms the common part of a broad class of adaptive algorithms.

The second step, time update $p(w(k) \,|\, H_j,\, k) \rightarrow p(w(k+1) \,|\, H_j,\, k)$, makes the estimator adaptive. In this paper we use for modeling of the time update step the standard exponential age-weighting. It corresponds to the operator

$$p(w(k+1) \,|\, H_j,\, k) \quad \propto \quad \{p(w(k) \,|\, H_j, k)\}^{\lambda}. \tag{5}$$

2.2. Normal distribution → recursive least squares

With the model (1) and the time update operator (5) chosen, we can derive the recursive formulae for the propagation of the conditional p.d.f. of unknown parameters (for details see e.g. [4]).

The computation is simplified if we assume the distribution of parameters $p(w(k)| \, H_j,\, k)$ normal. The general functional recursion is replaced by the algebraic recursion of parameter estimates and the covariance matrix $P(k)$. The linear part of the final algorithm reduces to the recursive least squares in this case. This can be computed in a square root factorised QR form [6] by the triangular Gentleman - Kung array Fig. 3. The array directly updates the upper triangular matrix $R(k)$ and a vector of parameters $\hat{r}(k)$ where $R'(k)R(k) = P^{-1}(k)$ and $\hat{r}(k) = R(k)\hat{w}(k)$.

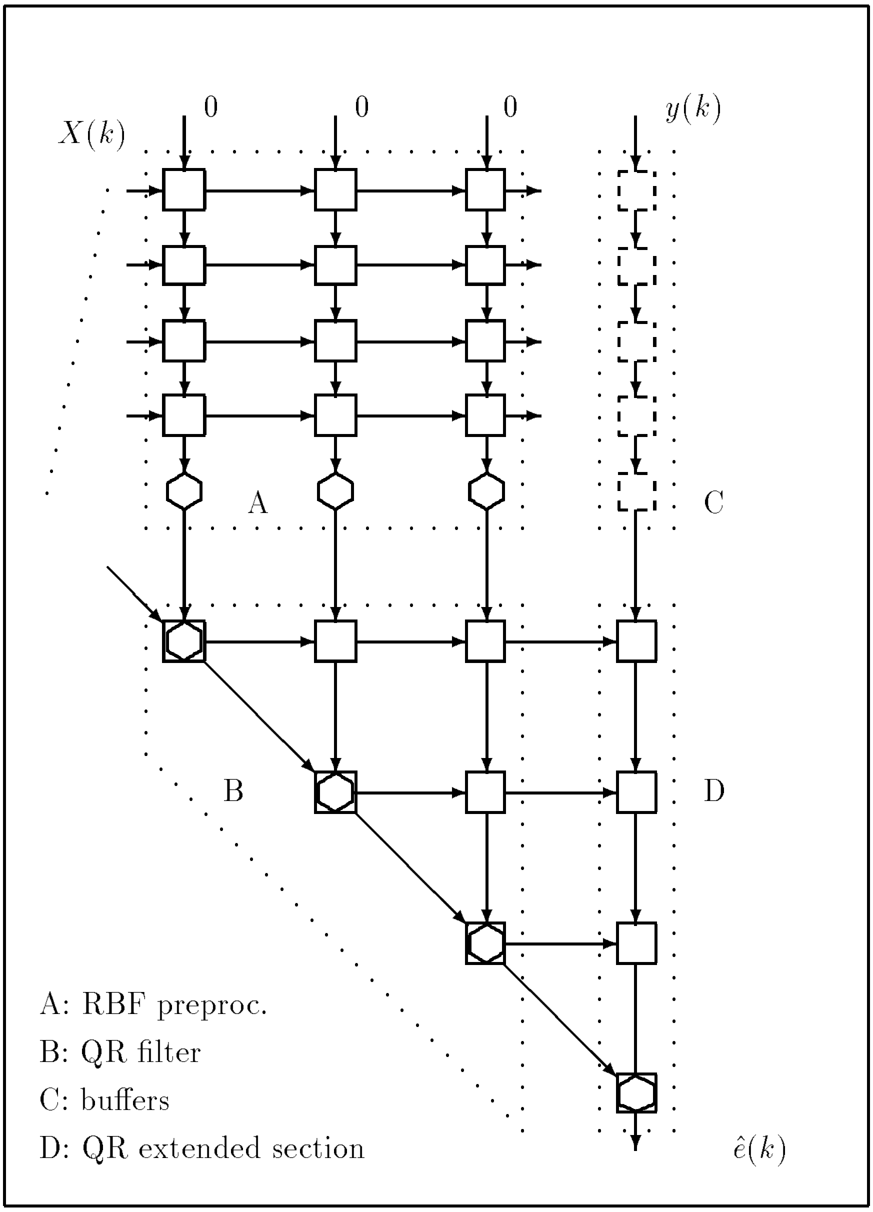

Figure 3. Radial basis function, systolic implementation.($m = 3$; $N_c = 4$)

Fig. 3. presents a combined array as proposed first by [5], for a known RBF values and structure. The upper triangular matrix $R(k)$ is updated in cells of processors of the section **B**. The section **D** updates vector of parameters $\hat{r}(k)$. The section **A** is a RBF nonlinear preprocessor analogical to Fig. 2.

3. RBF structure estimation concept

One way to incorporate our uncertainty about the exact structure of the Radial Basis Function model structure is to evaluate in parallel the probability of several models and, using this extra information, track and/or decide about the current optimal model from the processed data.

In this paper we propose, that we use $m-1$ fixed RBF's whose centres have been determined during the learning phase, plus one additional RBF which has a small set of N_H possible centres, characterising the parallel computed hypothesis $H_j, j = 1, 2, ...N_H$.

The recursive estimation of the probabilities assigned to the above hypotheses is based on Bayesian probability theory [4], [11].

The concept of the parallel computation is principally shown in Fig. 4. and Fig. 5. Because of a limited space, we can refer only about the basic concepts in this paper. The detailed derivations can be found in [11], [12] and [13].

3.1. System classification and Bayesian theory

The estimation of the probability of which centre is used for this additional RBF is one of the tasks, referred to as the classification of the system structure. As was already mentioned, in our case we have a set of the possible values of the additional RBF.

Our objective is to estimate and track in time N_H probabilities of alternative RBF's for the given input-output data observed on the system under the test. Using this information we can decide, which value from the set is the most probable one [4].

Let us denote $p(H_j(k)\,|\,k-1)$ the prior probability of the hypothesis $H_j(k)$, that the additional, (m-th) RBF of the system is equal to a j-th value from the updated set. The probability of the system output $y(k)$ (conditioned on the measured data and a given value of the tested RBF centre) is denoted $p(y(k)|\,H_j,\,k-1)$. Then, the aposterior probability of a hypothesis, conditioned on the measured data, is given by the Bayes formula.

$$p(H_j(k)\,|\,k) = \frac{p(y(k)|\,H_j,\,k-1)}{\sum_{i=1}^{N_H} p(y(k)|\,H_i,\,k-1)p(H_i(k)\,|\,k-1)} p(H_j(k)\,|\,k-1) \quad (6)$$

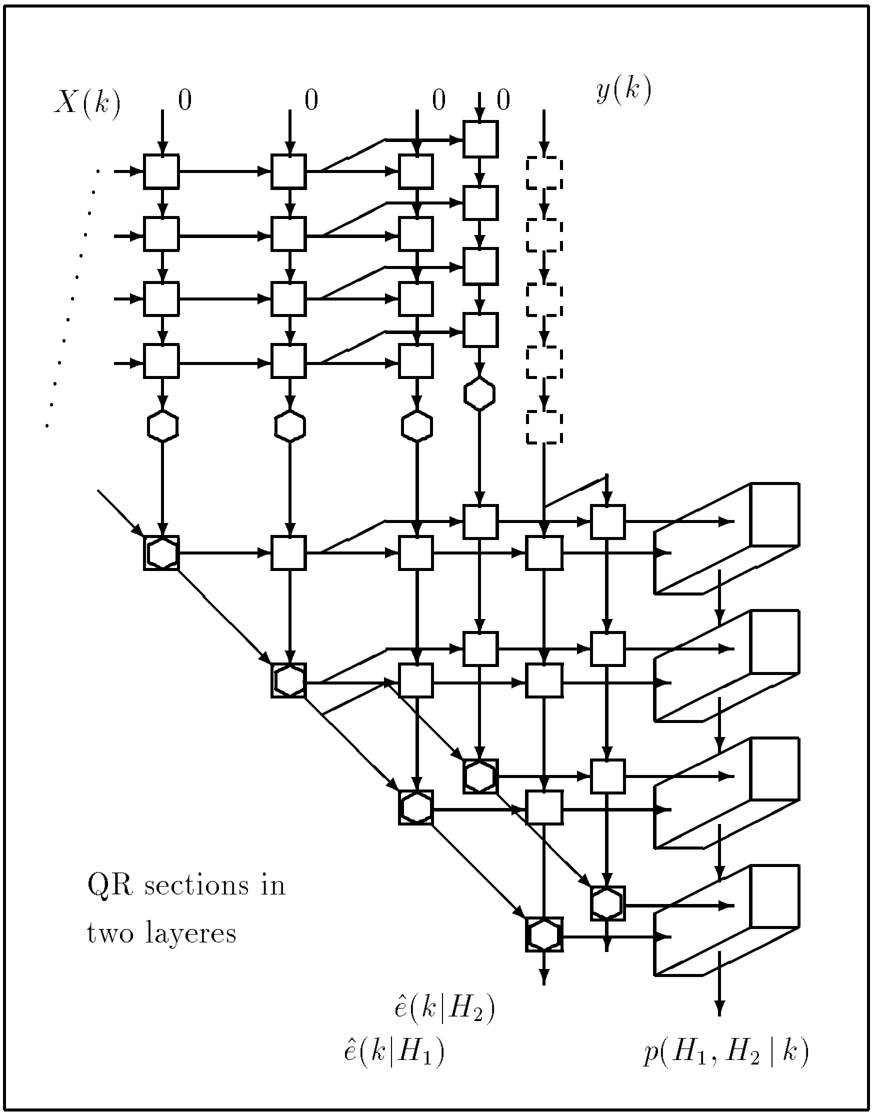

Figure 4. Systolic array for RBF structure estimation.
$(m = 3\,;\, N_c = 4;\, N_H = 2)$

The update of N_H probabilities is implemented in the last column of processors in Fig 4. The processors perform the recursive update of probabilities $p(H_j(k)\,|\,k-1) \rightarrow p(H_j(k)\,|k)$ by coordinates [4]. Processors compute the N_H scaling factors in (6) using the data stored and propagated in the arrays associated with each hypothesis H_j.

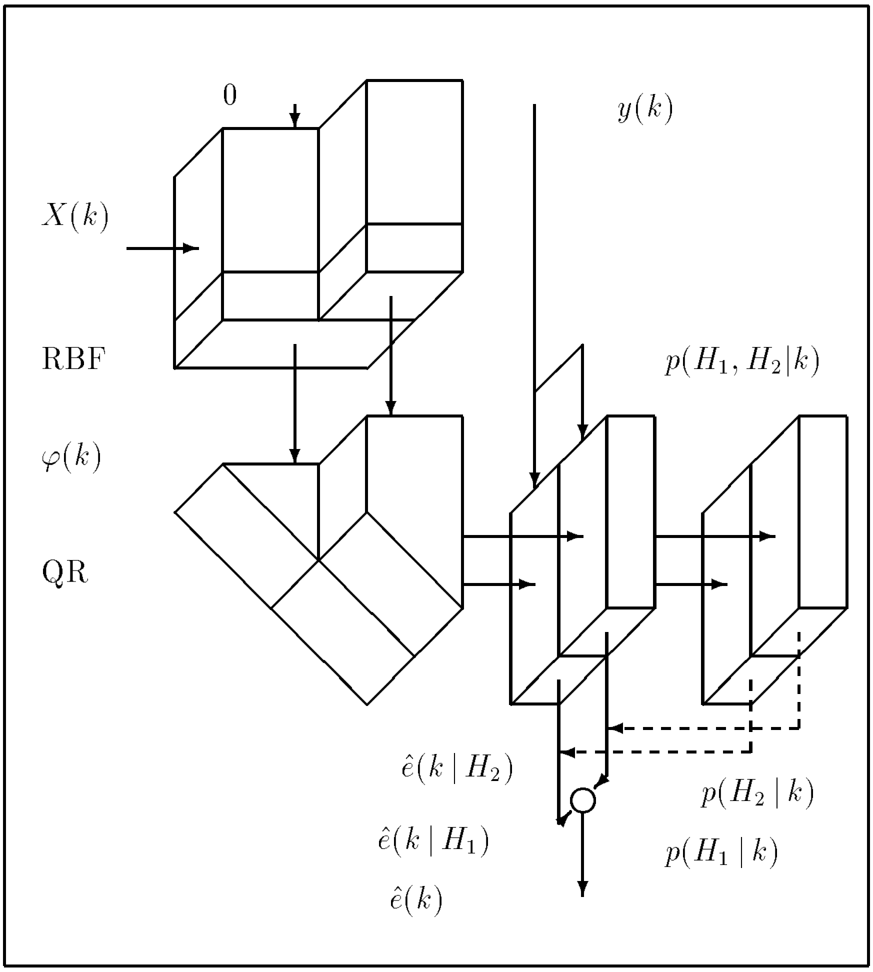

Figure 5. Systolic array for RBF structure estimation – block form.

Fig. 5. describes in a block form how the QR architecure can be modified, to incorporate the recursive update of probabilities of the hypothesis about the RBF structure (6) (7). Compare with Fig.3. and [5],[8] - [10].

Fig. 5. presents the part of the nested cells of the RBF and QR arrays shared by both models. The computations in the nested sections are identical for both compared models and therefore computed just once.

Estimated probabilities are used as the weights in an upper "meta" layer of the identification, constructing the optimal output prediction–error as a mixture of outputs from the alternative models.

In the practical applications (like the permanent monitoring, filtering, prediction or the event detection) the recursion (6) has to be complemented by the exponential weighting time update step in order to keep the tracking of the probabilities of the parallel updated RBF models

$$p(H_j(k+1)\,|\,k) \propto \{p(H_j(k)\,|\,k)\}^{\lambda_H}\,. \qquad (7)$$

In this way small changes in the system structure can be tracked by the recursive updates of the probabilities of the alternative RBF's.

The N_H probabilities (scalar numbers) determining the probability density function $p(H_j(k)\,|\,k)$ can be used to detect the change of the structure or operation mode of the estimated system.

4. Conclusion

The proposed array for the estimation of the RBF model can track small changes in the system by adapting the weights across all the RBF's.

It also recursively updates the probabilities of the hypothesis about the RBF structure. These probabilities (scalars) can be used to detect the change of the structure or operation mode of the estimated system.

The paper described the basic principals of

- a method for modelling of nonlinear systems based on RBF's and linear optimisation
- a method of incorporating different modes of operation into the model
- a concept of a neural network type architecture, based on the recursive QR decomposition algorithm to implement this method.

5. Acknowledgments

The author gratefully acknowledges support by the The International Fund for Ireland and The Royal Society UK-Czechoslovak Postdoctoral Fellowship Program.

References

[1] Powel, M. J. D.: Radial basis functions for multivariable interpolation: A review. in Proc. IMA Conf. on Algorithms for the Approximation of Functions and Data. RMCS Shrivenham, 1985.

[2] Rumelhart, D. E., G. E. Hinton, R. J. Williams Learning internal representations by Error propagation, in Parallel distributed processing. Volume 1 (MIT Press, 1987), pp. 318 - 362.

[3] Broomhead D. S. , Lowe D.: Multi – variable functional interpolation and adaptive networks", Complex Syst. 1988 2, pp. 321 - 355.

[4] Peterka, V. Bayesian approach to system identification In P. Eykhoff (Ed.) *Trends and Progress in System Identification*, Pergamon Press, Eindhoven, Netherlands, 1981.

[5] Mc Whirter, J.G., D. S. Broomhead, T. J. Shepherd: A Systolic array for nonlinear adaptive filtering and pattern recognition. In. Proceedings of Application Specific Array Processors, Princeton, New Jersey, 1990 IEEE Computer Society Press. Los Alamitos, Ca. pp. 700-711.

[6] Mc Whirter, J.G. Recursive Least Squares minimization using a Systolic Array, Proc. *SPIE*, Vol. 431, *Real Time Signal Processing VI*, pp. 105, 1983.

[7] Chen S., C. F. N. Cowan, P. M. Grant: Orthogonal least squares learning algorithm for Radial Basis Function Networks, IEEE Trams on Neural Networks, Vol 2 No 2. March 91.

[8] Mc Whirter, J.G. Algorithmic engineering - an emerging discipline, *SPIE* Vol. 1152, Advanced Algorithms and Architectures for Signal Processing IV, pp. 2-15, 1989.

[9] Gentleman,W.M., H.T.Kung. Matrix Triangularisation by Systolic Arrays. Proc. SPIE, Vol.298, *Real Time Signal Processing IV*, 1981, pp. 19-26.

[10] Gaston F. M. F., G. W. Irwin. A systolic square root information Kalman filter, Proc. IEEE Int. Conf. on Systolic Arrays, San Diego, California, May 1988, pp. 643-652, 1988.

[11] Krn M., 'Bayesian Estimation of Model Order', *Problems of Control and Information Theory*, Vol.**9** (1), 33–46 (1980).

[12] Kadlec, J. (1985). Recursive Bayesian Identification of Autoregression Model with Unknown Order Using Lattice Structures. *Report No. 1313*, Institute of Information Theory and Automation of the Czechoslovak Academy of Sciences, Prague (in Czech).

[13] Kadlec, J. (1986). Bayesian Identification of Regression Model in Fixed Point. *Ph.D. Thesis*, Institute of Information Theory and Automation of Czechoslovak Academy of Sciences Prague (in Czech).

[14] Gaston F. M. F., G. W. Irwin. The systolic approach to information Kalman filtering, *Int. Jour. Control*, Vol. 15., No. 1, pp. 225-228, July 1989.

[15] Gaston F.M.F. ,G. W. Irwin, J. G. Mc Whirter. Systolic Square Root Covariance Kalman Filtering, *Jour. of VLSI Sig. Processing* , No. 2, pp. 37-49, 1990.

[16] Gaston, F.M.F., G. W. Irwin, A Systolic parameter estimator for real-time adaptive control., Proc. IEE UK IT 90 Conference, 19-22 March 1990, conference publication 316, pp 111-116, 1990.

[17] Kadlec,J., F.M.F.Gaston, G.W.Irwin (1992). Parallel Implementation of Restricted Parameter Tracking. Chapter in: *Mathematics in Signal Processing* , J. McWhirter editor, Oxford Publishers Press.

[18] Kadlec, J., F.M.F. Gaston, G. W. Irwin, (1992) A Nonlinear Systolic Filter with On-line Radial Basis Function Estimation, In: Neural Computing Research and Applications (Edited by Adan Hilger), *Proceedings of the 2nd Irish Neural Network Conference*, Belfast, N. Ireland, 183-190.