Article ID: 1000-8152(2002)03-0381-06

Multidimensional Wavelet Networks Based on a Tensor Product Structure

WAN Jian, XU Demin and HE Yuyao
(College of Marine Engineering, Northwestern Polytechnical University·Xi'an, 710072, P.R. China)

Abstract: Based on the wavelet frame theory, a novel wavelet network for function learning in multidimensional spaces is proposed to avoid the 'curse of dimensionality'. The main feature of the proposed wavelet network is to multiply the reconstruction of each dimension in the output layer instead of adding them as usual. Thus a multidimensional wavelet frame will be generated automatically for approximation, and function learning can be realized through online or off-line adjustment of corresponding weight coefficients. Design methods for one-dimensional wavelet networks can also be generalized straightforwardly to multidimensional cases by using the tensor product structure. In the experiments, the multidimensional wavelet network performs well.

Key words: wavelet frames; multidimensional wavelets; tensor product structure; function approximation
Document code: A

基于张量积结构的多维小波网络

万　建　徐德民　贺昱曜
(西北工业大学航海工程学院·西安,710072)

摘要: 针对多维函数逼近的'维数灾'问题,依据小波框架理论提出了一种张量积结构小波网络,其主要特点是在网络输出层将各维输入的小波重构相乘,从而得到自动覆盖函数输入空间的多维小波框架,最后通过权系数的在线或离线学习实现多维函数的小波逼近.理论分析和仿真结果证实了该结构设计方法应用于多维函数逼近时的有效性.

关键词: 小波框架;多维小波;张量积结构; 函数逼近

1 Introduction

The wavelet transform, as a tool for signal analysis, was first proposed by J. Morlet in [1]. It soon emerged as a means of representing a function in a manner which readily reveals properties of the function in localized regions of the joint time-frequency space. Spatio-spectral properties of the wavelet transform provide a useful theoretical framework to investigate the structure of networks[2]. Thus the reconstruction of a function can be realized through linear combination of members of a preselected wavelet base or frame and weight coefficients of the combination can also be learned through computational architectures similar to neural networks. The idea of using wavelets in networks was first proposed by Q. Zhang and A. Benveniste[3] and Y. C. Pati and P. S. Krishnaprasad[4]. A systematic synthesis procedure for structural design of wavelet networks was also presented in their papers. However, those papers[5,6] focused on one-dimensional wavelet networks. The reason for it is that the implementation of wavelet networks of higher dimensions is of prohibitive cost if the same structure as in one-dimensional cases is used. For example, when single-scaling wavelets are used in separable form, as many as $2^n - 1$ mother wavelets are required, where n is the dimension of the wavelets[7]. Furthermore, it is hard to determine the lattice and number of dilation or translation parameters of multidimensional wavelet functions to suffice for generating a frame of $L_2(\mathbb{R}^n)$. Finally, the number of weight coefficients needed to adjust will increase sharply as to higher dimensions and thus the processes of learning will be too slow for practical applications. The problem incurred in

Received date: 2001-05-11; Revised dete: 2002-01-28.

multidimensional cases is also referred to as 'curse of dimensionality'.

In this paper we propose a novel multidimensional wavelet network structure for function learning on the basis of the wavelet frame theory[8]. The main idea of the proposed wavelet network is to multiply the reconstruction of each dimension in the output layer instead of adding them as usual. Thus a multidimensional wavelet frame can be generated automatically for approaching functions of interest. Design methods for one-dimensional wavelet networks are then generalized straightforwardly to multidimensional cases by using the tensor product structure. Learning algorithms for the multidimensional wavelet network and experimental results are given also.

2 Structural design

We will discuss systematic synthesis procedures for wavelet networks in this section. First, we briefly review the design methods for one-dimensional wavelet networks as analyzed in [3,4]. Then we generalize them to multidimensional cases by using the tensor product structure.

2.1 One-dimensional case

At first, Theorem 2.1 provides elementary insight for systematic design methods of one-dimensional wavelet networks. Its proof can be found in [9].

Theorem 1 Give a function $f \in L^2(\mathbb{R})$, which is mainly concentrated in $[-T, T]$ in time and whose Fourier transform $\hat{f}$ is concentrated mostly between the frequencies Ω_0 and Ω_1, namely, the signal f is essentially concentrated in the set $[-T,T] \times ([-\Omega_1, -\Omega_0] \cup [\Omega_0, \Omega_1])$. Suppose that integral dilations and translations of the mother wavelet $\varphi(x)$

$$\varphi_{mn}(x) = a_0^{-m/2}\varphi(a_0^{-m}x - nb_0) \tag{1}$$

constitute a frame, with frame bounds A, B, and dual frame $(\varphi_{mn})^{\sim}$. Assume that

$$|\hat{\varphi}(y)| \leqslant C|y|^{\beta}(1+y^2)^{-(\alpha+\beta)/2}, \tag{2}$$

where $C, \beta > 0, \alpha > 1$, and that, for some $\gamma > 1/2$

$$\int \mathrm{d}x(1+x^2)^{\gamma}|\varphi(x)|^2 < \infty. \tag{3}$$

Fix $T > 0, 0 < \Omega_0 < \Omega_1$. Then, for any $\varepsilon > 0$, there exists a finite subset $B_\varepsilon(T, \Omega_1, \Omega_2)$ of $\mathbb{Z}^2$ such that, for all $f \in L^2(\mathbb{R})$,

$$\| f - \sum_{(m,n)\in B_\varepsilon(T,\Omega_1,\Omega_2)} (\varphi_{mn})^{\sim}\langle \varphi_{mn}, f\rangle \| \leqslant$$

$$(B/A)^{1/2}[\|(I-Q_T)f\| + \|(I-P_{\Omega_1}+P_{\Omega_0}f\| + \varepsilon\|f\|], \tag{4}$$

where $(\Omega_T f)(t) = \chi_{[-T,T]}(t)f(t)$, $(P_\Omega f)^{\wedge}(w) = \chi_{[-n,n]}(w)\hat{f}(w)$, and χ_i denotes the indicator function of the interval I.

The theorem above demonstrates that any square integral function can be spanned by finite wavelet functions. Furthermore, it turned out to be essentially right that only those phase space lattice points, which were lying within the area $[-T, T] \times ([-\Omega_1, -\Omega_0] \cup [\Omega_0, \Omega_1])$, would suffice to approximately reconstruct f. So in practice, assuming that the time center and the frequency center of the selected mother wavelet $\varphi(x)$ are 0 and 1, respectively, then the set B_ε includes all pairs of (m, n) for which $\Omega_0 \leqslant a_0^{-m} \leqslant \Omega_1$, and $|a_0^m nb_0| \leqslant T$ if time-frequency focus areas of one dimensional functions can be estimated conveniently through the time-frequency analysis of observed data. Thus we can pre-determine the number and the concrete values of wavelet functions needed for approximating f, namely, the nodes or activation functions of wavelet networks can be determined definitely. On the other hand, the weight coefficients of the reconstruction can be learned through the computing structure shown in Fig. 1. Efficient algorithms of gradient-descent type can be adopted for learning because of the linearity of the networks.

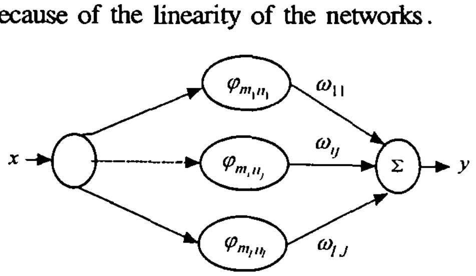


Fig. 1 One-dimensional wavelet networks

2.2 Multidimensional case

The synthesis methods in 2.1 encountered great challenge when applied directly to approximation of multidimensional functions. No matter what type of multidimensional wavelets is used, the number of multidimensional wavelet functions as well as the number of weight coefficients will increase sharply as to higher dimensions if the same structure in one-dimensional wavelet networks is used. Furthermore, because the estimation of spectral concentration of signals in high dimensions is computationally expensive, it becomes rather difficult to select the concrete lattice of those functions to cover the time-frequency domains of multidimensional functions of interest. There exist some cases that it is impossible or difficult to estimate the spectral concentration of signals, thus it is hard or unsuitable to pre-determine the dilation and translation parameters of wavelet functions acting as nodes of networks under such circumstances. Then dilation and translation parameters of multidimensional wavelet functions can be designed to be adjustable for determining their optimal values through another type of wavelet networks structure[3] by error back-propagation type algorithm. But this will incorporate much more additional parameters to be learned and exacerbate the 'curse of dimensionality'. In order to provide more insight into the novel multidimensional wavelet network structure proposed below, a lemma[8] is described in the following.

Lemma 1 Let $\Psi \in L^2(\mathbb{R}^n)$. For $a \in \mathbb{R}, a > 1, b = (b_1, \cdots, b_n) \in \mathbb{R}^n$ and $b_i > 0, i = 1, \cdots, n$, define the dilation and translation matrices D_j and T as $D_j = \operatorname{diag}(a^{j_1}, \cdots a^{j_n})$, where $j = (j_1, \cdots, j_n)^{\mathrm{T}} \in \mathbb{Z}^n$, and $T = \operatorname{diag}(b_1, \cdots, b_n)$, consider the family of translated and dilated functions of the form

$$\Psi(a,b) = \{\Psi_{j,k}(x) = \det D_j^{\frac{1}{2}} \Psi(D_j x - Tk) : j, k \in \mathbb{Z}^n\}, \tag{5}$$

if

$$m(\Psi, a) \triangleq \operatorname*{ess\,inf}_{|\omega_i| \in [1,a]} \sum_{j \in \mathbb{Z}^n} |\hat{\Psi}(D_{-j}\omega|^2 > 0, \tag{6}$$

$$m(\Psi, a) \triangleq \operatorname*{ess\,sup}_{|\omega_i| \in [1,a]} \sum_{j \in \mathbb{Z}^n} |\hat{\Psi}(D_{-j}\omega|^2 < \infty, \tag{7}$$

and

$$\sup_{\eta \in \mathbb{R}^n} [(1 + \eta^{\mathrm{T}}\eta)^{n(1+\varepsilon)/2} \beta(\eta)] = C_\varepsilon < \infty \tag{8}$$

for some $\varepsilon > 0$, where

$$\beta(\eta) \triangleq \sup_{|\omega_i| \in [1,a], i=1,\cdots,n} \sum_{j \in \mathbb{Z}^n} |\hat{\Psi}(D_{-j}\omega)| \cdot |\hat{\Psi}(D_{-j}\omega + \eta)|. \tag{9}$$

Then there exists $b_0 > 0$ such that $b_i \in (0, b_0), i = 1, \cdots, n$, the family defined above constitutes a frame for $L^2(\mathbb{R}^n)$; i.e., $\exists$ two constants $A > 0$ and $B < \infty$, such that $\forall f \in L^2(\mathbb{R}^n)$ the following inequalities hold

$$A \| f \|^2 \leqslant \sum_{j,k} |\langle \Psi_{j,k}, f \rangle|^2 \leqslant B \| f \|^2. \tag{10}$$

In practice, we are interested in a methodology that allows us to construct the multidimensional wavelet functions leading to frames, i.e., to find a mother wavelet function that satisfies, together with its dilation and translation parameters, sufficient conditions outlined in the above lemma. In the following, we discuss a tensor product construction of multiscaling wavelet frames.

Let $\Psi(x)$ be a tensor product of $1-D$ wavelet functions, i.e.,

$$\Psi(X) = \Psi_1(x_1) \cdots \Psi_n(x_n), \tag{11}$$

then

$$\hat{\Psi}(\omega) = \hat{\Psi}_1(\omega_1) \cdots \hat{\Psi}_n(\omega_n). \tag{12}$$

$\Psi_i(x_i), i = 1, \cdots, n$, must satisfy the admissibility condition

$$\int \frac{|\hat{\Psi}_i(\omega_i)|^2 \mathrm{d}\omega_i}{|\omega_i|} < \infty \tag{13}$$

under mild conditions of decay, this is satisfied if we choose $\Psi_i(x_i)$ such that

$$\int \Psi_i(x_i) \mathrm{d}x_i = 0. \tag{14}$$

If these $1-D$ functions can constitute frames, they must satisfy the first two conditions outlined in Lemma 1. These conditions are known to be necessary conditions as well in $1-D$[7]. The assumption made on mild decay conditions ensures that the second and third conditions are satisfied, and hence all conditions of Lemma 1 when reduced to $1-D$ are satisfied.

In the multidimensional case, by using the inequalities in $1-D$ above, and the fact that the infimum and supremum can now be taken over by the sum in each dimension, we have

$$m(\Psi, a) = \operatorname*{ess\,inf}_{|\omega_i| \in [1,a], i=1,\cdots,n} \{\sum_{j_1} |\hat{\Psi}_1(a^{-j_1}\omega_1|^2 \cdots \sum_{j_n} |\hat{\Psi}_n(a^{-j_n}\omega_n|^2\} > 0, \tag{15}$$

$$M(\Psi, a) = \operatorname*{ess\,sup}_{|\omega_i| \in [1,a], i=1,\cdots,n} \{\sum_{j_1} |\hat{\Psi}_1(a^{-j_1}\omega_1|^2 \cdots \sum_{j_n} |\hat{\Psi}_n(a^{-j_n}\omega_n|^2\} < \infty. \tag{16}$$

The third condition is to verify the convergence of the multi-indexed series[9]

$$\sum_{|k| \neq 0} [\beta(2\pi T^{-1}k)\beta(-2\pi T^{-1}k)]^{\frac{1}{2}}. \tag{17}$$

We have the following inequality

$$\sum_{|k|\neq 0}[\beta(2\pi T^{-1}k)\beta(-2\pi T^{-1}k)]^{1/2}\leqslant$$
$$\sum_{k_1}(1+(2\pi b_1^{-1}k_1)^2)^{-\frac{(1+\varepsilon)}{2}}\cdots\sum_{k_n}(1+(2\pi b_1^{-1}k_1)^2)^{-\frac{(1+\varepsilon)}{2}}\leqslant$$
$$\sum_{k_1}|2\pi b_1^{-1}k_1|^{-(1+\varepsilon)}\cdots\sum_{k_n}|2\pi b_1^{-1}k_n|^{-(1+\varepsilon)}. \tag{18}$$

Since each sum over k_i converges, $i=1,\cdots,n$, the sum involving β converges. Moreover, as $b_i\to 0, i=1,\cdots,n$, this sum tends to 0. Hence all conditions of Lemma 1 are satisfied. Therefore the tensor product of the reconstruction of each dimension leads to valid wavelet frames of multidimensional spaces. Having this in mind, we propose the novel multidimensional wavelet network as illustrated in Fig. 2. Its construction lies in two steps: first, similar structures to Fig. 1 are designed for every input using the same method in 2.1, namely, dilation and translation parameters of (m,n) are pre-selected to cover the whole spatio-spectral domain of each dimension when the spectral concentration for each dimension can be estimated without much difficulty. Otherwise, dilation and translation parameters of each dimension should be designed to be adjustable by using the structure proposed by Q. Zhang. Second, the reconstructions of each dimension are multiplied in the output layer to generate frames for multidimensional spaces automatically according to Lemma 1. Assume that there are N_i pre-selected wavelet functions for the ith input, where $i=1,\cdots,d$, the novel structure will generate as many as $\prod_{i=1}^{d}N_i$ multidimensional wavelet functions to cover the whole input space autonomously through such a tensor product. Weight coefficients as little as $\sum_{i=1}^{d}N_i$ need to be adjusted instead of $\prod_{i=1}^{d}N_i$ as usual. Thus the number of parameters needed to adjust is greatly decreased and the process of learning will be accelerated so as to avoid the 'curse of dimensionality'.

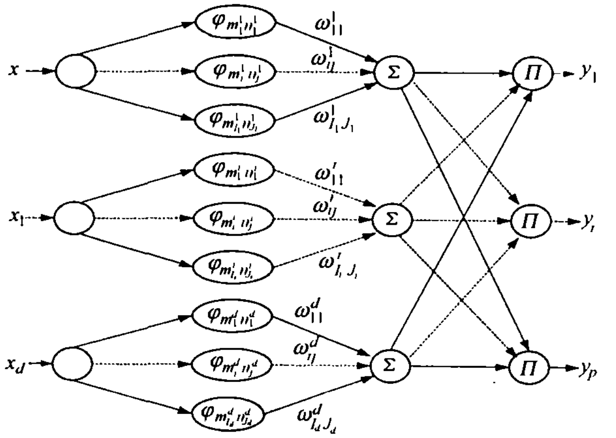


Fig. 2 Multidimensional wavelet networks

3 Learning algorithms

In this section the learning is based on a sample of random input/output pairs $\{X,F(X)\}$, where X is the input vector and $f(X)$ is the function to be approximated. For the sake of brevity, we assume that dilation and translation parameters of each dimension can be pre-determined and only weight coefficients in Fig. 2 need to be learned. The assumption will not lose generality because algorithms discussed below can be extended directly to the cases where dilation and translation parameters need to be adjustable for optimal values through error backpropagation. We develop an online learning algorithm of gradient type, which is similar to the backpropogation algorithm for usual neural networks. An offline learning algorithm of nonlinear LSE type is also discussed. More precisely, in the sequel we are given a sequence of random pairs $\{X_k, Y_{dk}=f(X_k)\}$, where X_k, Y_{dk} are the input and desired output vectors, respectively. In this paper, we limit our research to single output case, namely, Y_{dk} is abbreviated to scalar y_{dk}. All the results about single output cases can be generalized to multi-output cases straightforwardly.

Based on Fig. 2, multidimensional wavelet network approximant is expressed as follows

$$f_N(X)=\prod_{d=1}^{D}\sum_m\sum_n\omega_{mn}^{d}\varphi_{mn}^{3}(x_d), \tag{19}$$

where $\varphi_{mn}(x_d)=2^{-m/2}\varphi(a_0^{-m}x_d-nb_0)$, $X=[x_1,\cdots,x_D]^{\mathrm T}$ is the input vector and D is the dimension, $m,n\in\mathbb{Z}$.

3.1 Gradient algorithm

Collect all the parameters ω_{ij} in a vector W and write y_N to refer to the output of the network defined by (19). The gradient algorithm is to minimize the following objective function

$$E(W)=\frac{1}{2}E\{[y_d-y_N]^2\}. \tag{20}$$

We prefer to implement an online stochastic gradient algorithm to recursively minimize the criterion (20) using input/output observations. This algorithm modifies the parameter vector W after each measurement (X_k, y_{dk}) in the opposite direction of the gradient of the function

$$e(W) = \frac{1}{2}[y_{dk} - y_{Nk}]^2. \qquad (21)$$

The partial derivative of the function (21) with regards to ω_{mn}^i is

$$\frac{\partial e}{\partial \omega_{mn}^i} = \varphi_{mn}^i(x_i) \prod_{d=1, d\neq i}^{D} \sum_m \sum_n \omega_{mn}^d \varphi_{mn}^d(x_d), \qquad (22)$$

where $i = 1, \cdots, D$. Thus the iterative gradient-descent procedure is

$$\omega_{mn}^i(k+1) = \omega_{mn}^i(k) - \eta \frac{\partial e}{\partial \omega_{mn}^i}, \qquad (23)$$

where η is the learning coefficient.

3.2 Nonlinear LSE algorithm

The error criterion that is generally used in the off-line training of networks is the minimization of a sum of square error functions[10]. For the network (19) and after the presentation of S pairs of input and desired output patterns, this function is

$$E_S(W) = \sum_{s=1}^{S} [y_{ds} - y_{Ns}]^2. \qquad (24)$$

We define $F(W_k) = [f_1(W_k), \cdots, f_s(W_k)]^T$, where $f_s(W_k) = y_{ds} - y_{Ns}$, $A = \begin{bmatrix} \nabla f_1(W_k)^T \\ \vdots \\ \nabla f_s(W_k)^T \end{bmatrix}$, and k is the iterative step. One of the Marquardt-Levenberg algorithms for the optimization problem is as follows:

1) Set the original value of $W_1, \alpha > 0, \beta > 1$, approved error $\varepsilon > 0$, compute $E_S(W_1), \alpha_1 = \alpha; k = 1$;

2) Set $\alpha := \alpha/\beta$, compute $F(W_k), A_k$;

3) Solve the equation $(A_k^T A_k + \alpha I) d_k = -A_k^T F(W_k)$, and get $W_{k+1} = W_k + d_k$;

4) Compute $F(W_{k+1})$, if $F(W_{k+1}) < F(W_k)$, then go to 6), or else to 5);

5) If $\| A_k F(W_k) \| \leqslant \varepsilon$, then stop the iteration and get the solution $W = W_k$; or else, set $\alpha := \beta\alpha$ and go to 3);

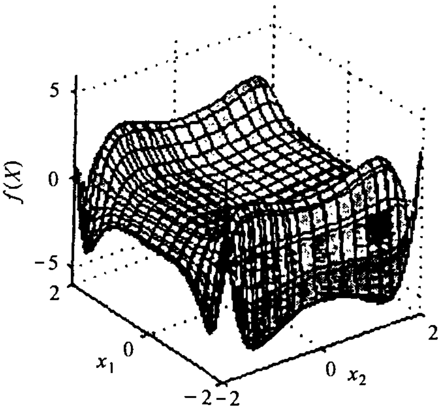
6) If $\| A_k F(W_k) \| \leqslant \varepsilon$, then stop the iteration and get the solution $W = W_k$; or else, set $k := k+1$ and go to 2).

4 Experimental results

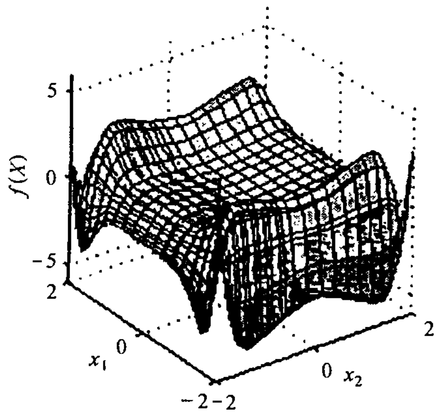
The multidimensional wavelet network described in Section 2.2 is tested in function learning, which is also compared with the MLP and Q. Zhang's wavelet network. Experiments have been performed for learning a two and three-dimensional function through an online and off-line algorithm, respectively.

First, we select a two-dimensional and square integral function in our off-line learning, in particular, to approximate the function $f(X) = (1 - \frac{\sin(4x_1)}{x_1})(1 - \frac{\sin(3x_2)}{x_2})$ over the limited time domain $l = [-2,0) \cup (0,2] \times [-2,0) \cup (0,2]$. Obviously, the frequency domain of the function is $[-4,4] \times [-3,3]$ according to its two-dimensional Fourier analysis. Thus we can determine the dilation and translation parameters for each dimension in advance by using the method described in Section 2.1, that is, for x_1 and x_2, we get the same parameters, i. e., $m \in \{-7, \cdots, 2\}$ and $n \in \{\text{floor}(-2^{m+1}) \cdots, \text{ceil}(2^{m+1})\}$ for each value of m, where floor and ceil mean to integrate toward $-\infty$ and $+\infty$, respectively. Now the computing structure of corresponding multidimensional wavelet network in Fig.2 has totally 104 weight coefficients to be learned. We select 961 pairs of (X, y_d) over l for the off-line learning of the function using the algorithm described in 3.2. Fig.3 illustrates the original form of $f(X)$ over l and its resulting approximation.

For the sake of comparability, another example is listed through a similar three-dimensional function and structure, which is $f(X) = [1 - \sin(4x_1)/x_1][1 - \sin(3x_2)/x_2][1 - \sin(2x_3)/x_3]$ over the limited time domain $\xi = [-2,0) \cup (0,2] \times [-2,0) \cup (0,2] \times [-2,0) \cup (0,2]$. The dilation and translation parameters of $1-D$ wavelet functions in the network are predetermined by using the method discussed above. Table 1 shows the approximation results obtained by three types of networks using online gradient-typed algorithms with the same learning coefficient $\eta = 0.01$. The networks are learned with 29791 random measurement points over ξ and the results are tested on the last 4000 points to compute square errors δ. Although we can only get a nonlinear-LSE-typed solution for the approximation, its resulting approximation is superior to the proposed wavelet network and MLP without rotation matrices. That is because the product structure of our wavelet network generates as many as 52 × 52 × 35 (i. e., 94640) 3 - D wavelet functions for approximation instead of 160 such 3 - D wavelet functions generated by dilating and translating the mother wavelet function $\varphi(X) = x_1x_2x_3e^{-(x_1^2+x_2^2+x_3^2)/2}$ in Q. Zhang's wavelet network.



(a) Original function



(b) Approximation result

Fig. 3 Approximation of the function $f(X)=(1-\frac{\sin(4x_1)}{x_1})(1-\frac{\sin(3x_2)}{x_2})$ over I

5 Conclusions

In this paper, a product-structured wavelet network for function learning in multidimensional spaces has been proposed. The novel structure is inspired by the principle of tensor product wavelet frames and is designed to avoid the 'curse of dimensionality'. The basic idea is to multiply the reconstruction of each dimension in the output layer instead of adding them as usual. The dilation and translation parameters of each dimension can also be designed to be adjustable through error backpropagation when the spatio-spectral domain of each input is hard to determine in advance. The multidimensional wavelet network proposed can greatly decrease the number of weight coefficients to be learned and shun the expensive work of selecting the lattice of multidimensional wavelet functions. The efficacy of such wavelet networks in multidimensional function learning is demonstrated through theoretical analysis and experimental results. As to the multidimensional wavelet network proposed, its essence is to transfer the linear optimization problem of an enormous number of parameters to the nonlinear optimization problem of a small number of parameters.

References

[1] Morlet J, Arens G, Fourgeau I, et al. Wave propagation and sampling theory [J]. Geophysics, 1982, 47(2):203 - 236

[2] Erdol N and Basbug F. Wavelet transform based adaptive filters: Analysis and new results [J]. IEEE Trans. Signal Processing, 1996, 44(9):2163 - 2171

[3] Zhang Q and Benveniste A. Wavelet networks [J]. IEEE Trans. Neural Networks, 1992, 3(6): 889 - 898

[4] Pati Y C and Krishnaprasad P S. Analysis and synthesis of feed-forward neural networks using discrete affine wavelet transformations [J]. IEEE Trans. Neural Networks, 1993, 4(1):73 - 85

[5] Zhang J, Walter G, Miao Y, et al. Wavelet neural networks for function learning [J]. IEEE Trans. Signal Processing, 1995, 43 (6): 1485 - 1497

[6] Delyon B, Juditsky A and Benvenviste A. Accuracy analysis for wavelet approximation [J]. IEEE Trans. Neural Networks, 1995, 6 (2): 332 - 348

[7] Daubechies I. Ten Lectures on Wavelets [M]. Philadelphea, PA, USA: SIAM Press, 1992

[8] Kugarajah T and Zhang Q. Multidimensional wavelet frames [J]. IEEE Trans. Neural Networks, 1995, 6(6): 1552 - 1556

[9] Daubechies I. The wavelet transform, time-frequency localization and signal analysis [J]. IEEE Trans. Information Theory, 1990, 36(5): 961 - 1005

[10] Kollias S and Anastassiou D. An adaptive least squares algorithm for the efficient training of artificial neural networks [J]. IEEE Trans. Circuits and Systems, 1989, 36(8): 1092 - 1101

本文作者简介

万　建　1975 年生.博士研究生.研究方向为自适应控制,神经网络与小波分析.

徐德民　1937 年生.教授,博士生导师.目前主要研究方向为非线性控制,智能控制和精确制导等.　Email: xudm@nwpu.edu.cn

贺昱曜　1956 年生.教授,博士.主要研究方向为神经网络,非线性控制以及现代电力电子技术等.