Front. Electr. Electron. Eng. China 2008, 3(2): 198–203
DOI 10.1007/s11460-008-0026-2

RESEARCH ARTICLE

Shuyuan YANG, Licheng JIAO, Min WANG

A new directional multi-resolution ridgelet network

Abstract In this paper, we propose a new directional multi-resolution ridgelet network (DMRN) based on the ridgelet frame theory, which uses the ridgelet as the activation function in a hidden layer. For the multi-resolution properties of the ridgelet function in the direction besides scale and position, DMRN has great capabilities in catching essential features of direction-rich data. It proves to be able to approximate any multivariate function in a more stable and efficient way, and optimal in approximating functions with spatial inhomogeneities. Besides, using binary ridgelet frame as the mathematical foundation in its construction, DMRN is more flexible with a simple structure. The construction and the learning algorithm of DMRN are given. Its approximation capacity and approximation rate are also analyzed in detail. Possibilities of applications to regression and recognition are included to demonstrate its superiority to other methods and feasibility in practice. Both theory analysis and simulation results prove its high efficiency.

Keywords ridgelet frame, directional multi-resolution, ridgelet network

1 Introduction

The mimicking of the structure and function of human brains by using artificial neurons is an improving process through the dozens of years of development of feed-forward neural network (FNN). The initial M-P model uses the Sigmoid function as the activation function in the hidden layer is a case in point. However, its infinite support violates the local characteristic of human neurons, which may require a lot of neurons in order to approximate an unknown function [1]. The radical basis function neural network (RBFNN) based on the regularization theory has been widely adopted in recent years. It adopts the radical basis function with the local characteristic as its activation function in the hidden layer, such as the Gaussian function, which improves the performance of the network remarkably [2]. Nonetheless, in RBFNN the basis functions in the hidden layer are redundant. The later developed wavelet neural network (WNN) introduces the wavelets function into the network to serve as the transform functions in the hidden layer [3]. For wavelets have the local characteristic both in time and frequency, WNN can mimic the multi-resolution properties of visual neurons and the multi-channel feature of hearing neurons. Compared with other FNN models, WNN has relative higher efficiency in solving practical problems.

For the activation function of neurons wavelet is a good choice but with insufficient flexibility of directions in high dimensional spaces. Recent research results of some psychologists show that the reception field of visual neurons of human being has the multi-scale characteristics in direction besides the scale and position [4]. In other words, there are some specified neurons that take charge of some specific directions, which will give the strongest response in some directions. It is the reason why the human neurons are sensitive to the directions, and a good mimic of this phenomenon can improve the accuracy of neural network models.

2 Ridgelet frame

In 1996, a pioneering work was done by Candes to represent multivariate functions with a new basis function in a more stable and efficient way [5]. Thus a new tool in harmonic analysis, ridgelet is proposed using the Littlewood-Paley theory. Its definition is as follows:

If $\psi: R^d \to R$ satisfies the condition:

$$K_\psi = \int \frac{|\hat{\psi}(\xi)|^2}{|\xi|^d} d\xi < \infty, \quad (1)$$

where $\hat{\psi}(\cdot)$ is the Fourier transform of the function ψ.

Translated from *Journal of Xidian University*, 2006, 33(4): 557–562
[译自: 西安电子科技大学学报(自然科学版)]

Shuyuan YANG (✉), Licheng JIAO
Institute of Intelligent Information Processing, Department of Electrical Engineering, Xidian University, Xi'an 710071, China
E-mail: syyang@xidian.edu.cn

Min WANG
Key Laboratory of Radar Signal Processing, Department of Electrical Engineering, Xidian University, Xi'an 710071, China
E-mail: wangmin@xidian.edu.cn

Then we call the function ψ as the admissible activation neuron function, and the functions $\psi_\gamma(\boldsymbol{x}) = a^{-1/2}\psi(\boldsymbol{u}\cdot\boldsymbol{x} - b)/a$ derived from the function ψ as ridgelets. The parameter $\gamma = (a,\boldsymbol{u},b)$ belongs to a space of neurons:

$$\Gamma = \{\gamma = (a,\boldsymbol{u},b), a,b \in R, a > 0, \boldsymbol{u} \in S^{d-1}, \|\boldsymbol{u}\| = 1\},$$

where $a,\boldsymbol{u},b$ represent the scale, the direction and the localization of ridgelet respectively. The cross section of the ridgelet is a wavelet-like curve, while along the ridge there is a line. It is this kind of geometric structure that makes the ridgelet deal with the directional information more efficiently. Discretizing γ, we can obtain such a group of discrete ridgelets:

$$\psi_\gamma(\boldsymbol{x}) = a_0^{j/2}\psi\left(a_0^j\boldsymbol{u}\boldsymbol{x} - kb_0\right),$$

where $\gamma \in \Gamma_d = \left\{\left(a_0^j,\boldsymbol{u},kb_0a_0^j\right), j \geqslant 0, \boldsymbol{u} \in \sum_j, k \in Z\right\}$. Here the scale parameter a is $\left\{a_0^j\right\}_{j \geqslant j_0}$ ($a_0 > 1$ and j_0 is the biggest scale); the localization parameter b is $\left\{kb_0a_0^{-j}\right\}_{k,\, j \geqslant j_0}$. At the scale a_0^j, the discrete sampling set on the sphere is denoted by $\sum_j$, which is ε_j-spaced grid on the sphere S^{d-1} (for $\varepsilon_0 > 0, \varepsilon_j = \varepsilon_0 a_0^{-(j-j_0)}$) in the d-dimensional space. Then γ is determined by the scale j. The finer the scale, the denser the sampling on the sphere S^{d-1}. The directional vector $\boldsymbol{u}$ in d-dimension is [6]:

$$u_1 = \cos\theta_1, u_2 = \sin\theta_1\cos\theta_2, \cdots, u_d = \sin\theta_1\sin\theta_2\cdots\sin\theta_d,$$
$$0 \leqslant \theta_1,\cdots,\theta_{d-1} \leqslant \pi.$$

A commonly used selection is the binary frame with $a_0 = 2$, $b_0 = 1$, $j_0 = 0$:

$$\left\{\varphi(u_i\boldsymbol{x} - k),\, 2^{j/2}\psi(2^j u_i^j\boldsymbol{x} - k),\, j \geqslant 0,\, u_i^j \in \sum_j,\, k \in Z\right\}, \quad (2)$$

where the resolution of $\sum_j$ is $\theta_0 2^{-j}$. The binary ridgelet frame in $2-d$ dimension is:

$$\left\{\varphi\left(x_1\cos\theta_i + x_2\sin\theta_i - k\right),\, 2^{j/2}\psi\left(2^j\left(x_1\cos\theta_i^j + x_2\sin\theta_i^j\right) - k\right),\right.$$
$$\left. j \geqslant 0,\, \theta_i^j = 2\pi\theta_0 i 2^{-j},\, k \in Z\right\}, \quad (3)$$

when $d = 2$. At scale j, θ_i^j is a set of points with equal space $2\pi\theta_0 2^{-j}$ on a circle (where θ_0^{-1} is an integer). The training of a neural network can be considered as the approximation of a group of input-output samples. Take the scale function φ and ridgelet function ψ as the activation functions of a three-layer FNN; and define the coefficients c and d as their corresponding connection weights, then we can gain the equation of the ridgelet network based on the discrete frame as

$$\hat{y} = f_0(\boldsymbol{x}) + \sum_{j=0}^{J} g_j(\boldsymbol{x})$$
$$= \sum_{i=0}^{\theta_0^{-1}/2}\sum_k c_{i,0,k}\varphi_{i,0.k}(\boldsymbol{x})$$
$$+ \sum_{j=0}^{J}\sum_{i=0}^{\theta_0^{-1}2^j}\sum_k d_{i,j,k}\psi_{i,j,k}(\boldsymbol{x}). \quad (4)$$

3 Directional multi-resolution ridgelet network (DMRN)

3.1 The network model and the determination of the hidden neurons

Base on Eq. (4), we get the model of ridgelet network shown in Fig. 1.

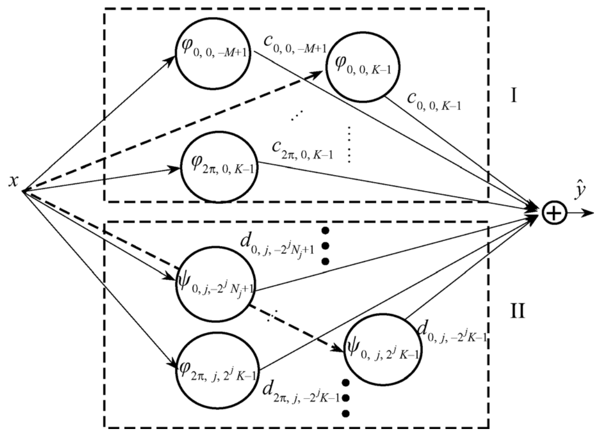

Fig. 1 Structure of the ridgelet network

It consists of two sub networks, I and II formed by the scale function $\varphi_{i,0,k}$ and the ridgelet function $\psi_{i,j,k}$ respectively, denoted as φ_0 (Support(φ_0) = $[0,M]$) and ψ_j (Support($\psi_j)_d = [0,N_j]$). Suppose the tight support of a d-dimension function f is Support(f) $= [0,L_1] \times [0,L_2] \times \cdots \times [0,L_d]$. Let $\sum_{i=1}^{d} L_i = K$. Equation (4) can be rewritten as follows:

$$\hat{y} = \sum_{i=0}^{\theta_0^{-1}/2}\sum_{k=-M+1}^{K-1} c_{i,0,k}\varphi_{i,0.k}(\boldsymbol{x})$$
$$+ \sum_{j=0}^{J}\sum_{i=0}^{\theta_0^{-1}2^j}\sum_{k=-2^jN_j+1}^{2^jK-1} d_{i,j,k}\psi_{i,j,k}(\boldsymbol{x}). \quad (5)$$

Equation (5) gives the network equation derived from a discretization of Γ. The discretization of the scale and the position are the same as that of WNN [3]. In this section the explanation focuses on the direction of discretization, that is, the important discretized u. From above we can see that in the points set corresponding to u, the distance of two adjacent points is of rank 2^{-j} [7,8]. If $\sum_j$ is of uniform distribution for any $j \geqslant j_0$, it can be deduced that $\sum_j$ is the set of L_j equal-spaced points on S^{d-1}, and L_j is of rank $2^{-j(d-1)}$. Through computation we obtain $L_j - L_{j+1} = O(2^{-j(d-1)}(2^{(d-1)}-1))$. Therefore, if the number of ridgelet directions in scale j has been determined, the number in scale $(j+1)$ can be deduced in such a way. First, divide the line connected by two adjacent points in scale j into ${a_0}^{(d-1)}$ portions and some points are hence obtained. Then at scale $(j+1)$ insert new directions at the corresponding positions of these points except two extreme points. These new points, together with the points in scale j are used to form the directions in scale $(j+1)$. Thus the new directions in scale $(j+1)$ are obtained.

Take the scale function and the ridgelet function as the activation functions in the hidden layer of the sub-networks I and II respectively. We can gain the following two lemmas based on the binary ridgelet frame theory, which lay some theoretical foundations for the determination of the number of neurons in the hidden layer of DMRN.

Lemma 1 The number of hidden neurons in the network I of DMRN is no more than $K+M-1$.

Proof The tight support of f is Support(f) = $[0,L_1] \times [0,L_2] \times \cdots \times [0,L_d]$. Thus the support of $\boldsymbol{u}\cdot\boldsymbol{x}$ is $[0,K]$. If $\boldsymbol{u}\cdot\boldsymbol{x}$ is taken as a new variable, then $\varphi_{i,0,k}$ can be regarded as having lost a dimension and thereafter becoming $\varphi_{0,k}$. Only the supports of these functions $\{\varphi_{0,-M+1}, \varphi_{0,-M+2},\cdots,\varphi_{0,0},\varphi_{0,1},\cdots,\varphi_{0,K-1}\}$ have nonempty intersections with $[0,K]$. That is, for $f_0(x) = \sum_k c_{0,k}\varphi_{0,k}$, only the above $K+M-1$ functions make contribution to $f_0 : f_0(x) = \sum_{k=-M+1}^{K-1} c_{0,k}\varphi_{0,k}$. Hence the number of neurons is no more than $K+M-1$.

Lemma 2 The number of hidden neurons in network II is no more than 2^jK+2^jN-1.

Proof Similar to Lemma 1, if $\boldsymbol{u}\cdot\boldsymbol{x}$ can be considered as a new independent variable, then $\psi_{i,j,k}$ can be considered as having lost a dimension and becoming $\psi_{j,k}$. So only the supports of the functions $\{\psi_{j,-2^jN_j+1}, \psi_{j,-2^jN_j+2},\cdots, \psi_{j,0}, \psi_{j,1},\cdots,\psi_{j,-2^jk-1}\}$ have nonempty intersections with $[0,K]$. That is, only the above functions make contribution to $g_j(x) = \sum_j\sum_k d_{j,k}\psi_{j,k}$, which can be written as $g_j(x) = \sum_{k=-2^jN_j+1}^{2^jL-1} d_{j,k}\psi_{j,k}$. Thus the number of the hidden neurons in the network II is no more than 2^jK+2^jN-1.

3.2 Training algorithm of the network

Suppose there are P training samples $[x_1,\cdots,x_p]$, $\hat{y}_p$ and y_p are the actual output and the expected output of the pth training sample x_p. Define such an error function E:

$$E = \|e\|^2 = \sum_p (\hat{y}_p - y_p)^2 = \sum_p \left[\sum_{l=0}^{\theta_0^{-1}/2}\sum_{m=-M+1}^{K-1} c_{i,0,k}\varphi_{i,0,k}(x_p) + \sum_{j=0}^{J}\sum_{i=0}^{\theta_0^{-1}2^j}\sum_{k=-2^jN_j+1}^{2^jK-1} d_{i,j,k}\psi_{i,j,k}(x_p) - y_p\right]^2. \tag{6}$$

The goal of training the network is to minimize E, so it can be reduced to an optimization problem. We can solve it using the gradient descending method. Using Eqs. (5) and (6), we can calculate the gradients of the coefficients:

$$\frac{\partial E}{\partial c_{i,0,k}} = 2\sum_p \left[\sum_{l=0}^{\theta_0^{-1}/2}\sum_{m=-M+1}^{K-1} c_{l,0,m}\varphi_{l,0,m}(x_p) + \sum_{n=0}^{j}\sum_{l=0}^{\theta_0^{-1}2^j}\sum_{m=-2^nN_n+1}^{2^nK-1} d_{l,n,m}\psi_{l,n,m}(x_p) - y_p\right] \times \varphi_{i,0,k}(x_p), \tag{7}$$

$$\frac{\partial E}{\partial d_{i,j,k}} = 2\sum_p \left[\sum_{l=0}^{\theta_0^{-1}/2}\sum_{m=-M+1}^{K-1} c_{l,0,m}\varphi_{l,0,m}(x_p) + \sum_{n=0}^{j}\sum_{l=0}^{\theta_0^{-1}2^j}\sum_{m=-2^nN_n+1}^{2^nK-1} d_{l,n,m}\psi_{l,n,m}(x_p) - y_p\right] \times \psi_{i,j,k}(x_p). \tag{8}$$

The updating equation of coefficients c and d are depicted as follows:

$$c_{i,0,k}(t+1) = c_{i,0,k}(t) - \eta \times \frac{\partial E}{\partial c_{i,0,k}},$$
$$i = 0,\cdots,\theta_0^{-1}/2; k = -M+1,\cdots,K-1, \tag{9}$$

$$d_{i,j,k}(t+1) = d_{i,j,k}(t) - \eta \times \frac{\partial E}{\partial d_{i,j,k}},$$
$$i = 0,\cdots,\theta_0^{-1}2^j; j = 0,\cdots,J; k = -M+1,\cdots,K-1. \tag{10}$$

Thus the procedure of the learning algorithm is then expounded.

Setp 1 Given the number of neurons in the hidden layer in the network I and the parameter θ_0, initialize the connected weights of the network.

Setp 2 Input the training samples and train the subnet I using Eq. (9), the stop criterion is

$$|E(t+1) - E(t)| < \varepsilon_1 \text{ or } |E(t)| < \varepsilon_2.$$

Here $\varepsilon_1,\varepsilon_2$ are two very small positive numbers.

Setp 3 Let $j = 0$, incorporate the network II into the network and initialize the weights and the number of neurons in the hidden layer.

Setp 4 Train the network with $c_{i,0,k}$ and $d_{i,j,k}$ ($l = 0,1,\cdots, j-1$) remaining unchanged; use Eq. (10) to learn the connected weights, the stop criterion is the same as that of Step 2.

Setp 5 Let $j = j+1$, when $j < J$ (a given maximum number), go to Step 4.

3.3 Properties of DMRN

The frame is a system from which we can get high-quality non-linear approximation [9]. The proposed DMRN is based on the binary ridgelet frame and the ridgelet theory [10]. As mentioned above, $\left\{\varphi(u_i x-k),\, 2^{j/2}\psi\left(2^j u_i^j x-k\right),\, j\geqslant j_0,\, u_i^j\in \sum_j,\, k\in Z\right\}$ is a binary frame in $L^2(R^d)$, and the scale function $\varphi(u_i\mathbf{x} - k)$ can be taken as a ridgelet function ψ when $j = -1$. Accordingly, the frame can be written as the superposition of the ridgelet functions ψ_γ:

$$D_{\Gamma_d} = \left\{\psi_\gamma,\, \gamma\in\Gamma_d\right\}.$$

In Ref. [10] Candés proved that this frame has quantitative bounds, and the ratio of the upper bound to the low bound can be determined by the ridgelet parameters $C(\theta_0+b_0)/(2\pi)$ [10].

Theorem 1 Let $\{\psi_\gamma\}$ be the ridgelet frame. Given a function f, and if for any $\xi\in R$, (φ,ψ) satisfies:

$$|\hat{\varphi}(\xi)|^2+\sum_{j>0}2^{j(d-1)}\left|\hat{\psi}(2^{-j}\xi)\right|^2=|\xi|^{d-1}. \qquad (11)$$

Then we can obtain

$$\left|\sum_\gamma |\langle f,\psi_\gamma\rangle|^2-\frac{1}{2\pi b_0\theta_0}\left\|\hat{f}\right\|_2^2\right|\leq C\left(\theta_0^{-1}+b_0^{-1}\right)\left\|\hat{f}\right\|_2^2. \qquad (12)$$

The frame ratio is bounded by $C(\theta_0+b_0)/(2\pi)$, where constant C at most relies on φ and ψ.

From Theorem 1 several conclusions can be obtained. First, the existence of bounds indicates that the frame has a proper discrete Γ in some sense. Second, the larger the ratio of the upper bound to the low bound is, the fewer items will appear in the expansion. Thus we can adjust the parameters in ridgelets to acquire a more efficient structure.

Theorem 2 The dictionary produced by the ridgelet frame can be used to approximate multivariate functions. Let f_M be an approximation by M items of the ridgelet function in the ridgelet frame:

$$f_M=\sum_{i=1}^{M}\lambda_{i,M}\psi_{\gamma_i,M}. \qquad (13)$$

If the ridgelet coefficients belong to the Lorentz weak l^p space $l_{p,\infty}$ and satisfy the condition $r = (1/p - 1/2)_+$, this approximation can achieve the approximation rate $\|f-f_M\|_2\leqslant CM^{-r}$ where C is a constant.

The Ref. [10] proves that D_{Γ_d} has the optimal approximation efficiency and indicates that for functions in the class $R^s_{p,q}(C)$, the representation with ridgelet frame is optimal. So the proposed DMRN constructed by ridgelet frame also has the optimal approximation in this sense, that is, the ridgelet network with superposition of ridgelet has optimal property. As long as the superposition coefficients of the ridgelet functions are correctly found, the network based on the frame theory has an optimal approximation for a large class of functions.

Finally we consider the approximation rate of the network. It proves that the approximation rate with the ridgelet is larger than Fourier transform (FT), Multilayer perceptron (MLP) and WNN for the functions above [10].

4 Simulation experiments and conclusions

Experiment 1 Consider 2-D singular functions f_1–f_3 defined on $[-1,1]^2$

$$f_1(x_1,x_2)=\frac{1_{\{10x_1-x_2>0\}}\sin(\pi y)}{2+\sin(2\pi x)},$$

$$f_2(x_1,x_2)=1_{\{x_1+4x_2>2\}}\exp\left(-x_1^2-x_2^2\right),$$

$$f_3(x_1,x_2)=1_{\{x_1-x_2>0\}}[1-4\sin c(4x_1)][1-3\sin c(3x_2)],$$

here f_1 is a function with linear singularity at $10x_1+x_2=0$, f_2 is a function with linear singularities at $x_1+4x_2=2$, f_3 is a function with singularities at $x_1-x_2=0$. Take 16 × 16 uniformly distributed samples as the training samples. In our proposed DMRN, let $J=5$, $\theta_0=0.1$, $\varepsilon_1=1\mathrm{e}-15$, $\varepsilon_2=1\mathrm{e}-20$. Let the learning steps of the DMRN described above and the WNN in Ref. [3] both be 0.01, and the stop condition of the two networks the same, as shown in the procedure in Sect. 2.2. The approximation results of DMRN and WNN are given in Table 1, where the approximated root mean square errors are given with different forms of ψ. From the comparison of the networks, for these functions with linear singularities, the proposed DMRN performs better than WNN.

Experiment 2 Classification is important in pattern recognition. Consider a three-class problem shown in Fig. 2(a). The whole region is divided by several straight lines, and the sub-regions that are classified are denoted by the circle, rectangle and bar respectively.

Because classification can be regarded as an approximation of a singular function with discrete output and linear singularities, it is more effective to use DMRN in solving the problem, compared with two other commonly used classification methods—RBF and support vector machine (SVM). Given 289 training samples, the Gaussian basis function is adopted in our proposed DMRN and RBF neural network. The parameters are the same as those in

Table 1 Approximation of singular function

	WNN			DMRN		
	f_1	f_2	f_3	f_1	f_2	f_3
Meyer	0.0091	0.0011	0.9940	3.2126e–004	1.6641e–004	0.0209
Haar	1.0401e–014	4.4322e–015	5.2592e–013	4.2652e–015	2.0226e–015	4.1164e–014
Db2	2.3669e–011	8.1038e–011	2.4744e–010	1.2401e–011	1.7930e–012	4.0330e–011
Db4	4.0862e–010	1.5219e–011	4.0908e–009	2.5134e–011	1.0005e–011	2.2615e–010
Coif1	4.6264e–010	1.3077e–011	5.0159e–009	2.8974e–011	7.8861e–012	2.2482e–010

Experiment 1, as shown in Figs. 2 (b–d) (the Gaussian basis function is employed in the networks). From the result we can see that DMRN can conduct a proper division of region with very small errors at the bottom, which is better than RBF and SVM.

The geometrical multi-resolution analysis (GMA) for high dimensional data has experienced a rapid development in recent years. As an extension of wavelet to a higher dimension [11], ridgelet has been a research hotspot [6–16]. A directional vector is defined in the ridgelet function, so the ridgelet has the multi-resolution properties in the direction besides scale and position, which accords with the characteristic of real visual neurons of human beings. In this paper, a three-layer directional multi-resolution ridgelet network (DMRN) is proposed based on the adoption of ridgelet as the activation function in the hidden layer, then a binary ridgelet frame is used for its construction.

First, for the good characteristic of the ridgelet, this new network can deal with data with hyperplane singularity efficiently as well as data in higher dimension. Second, the

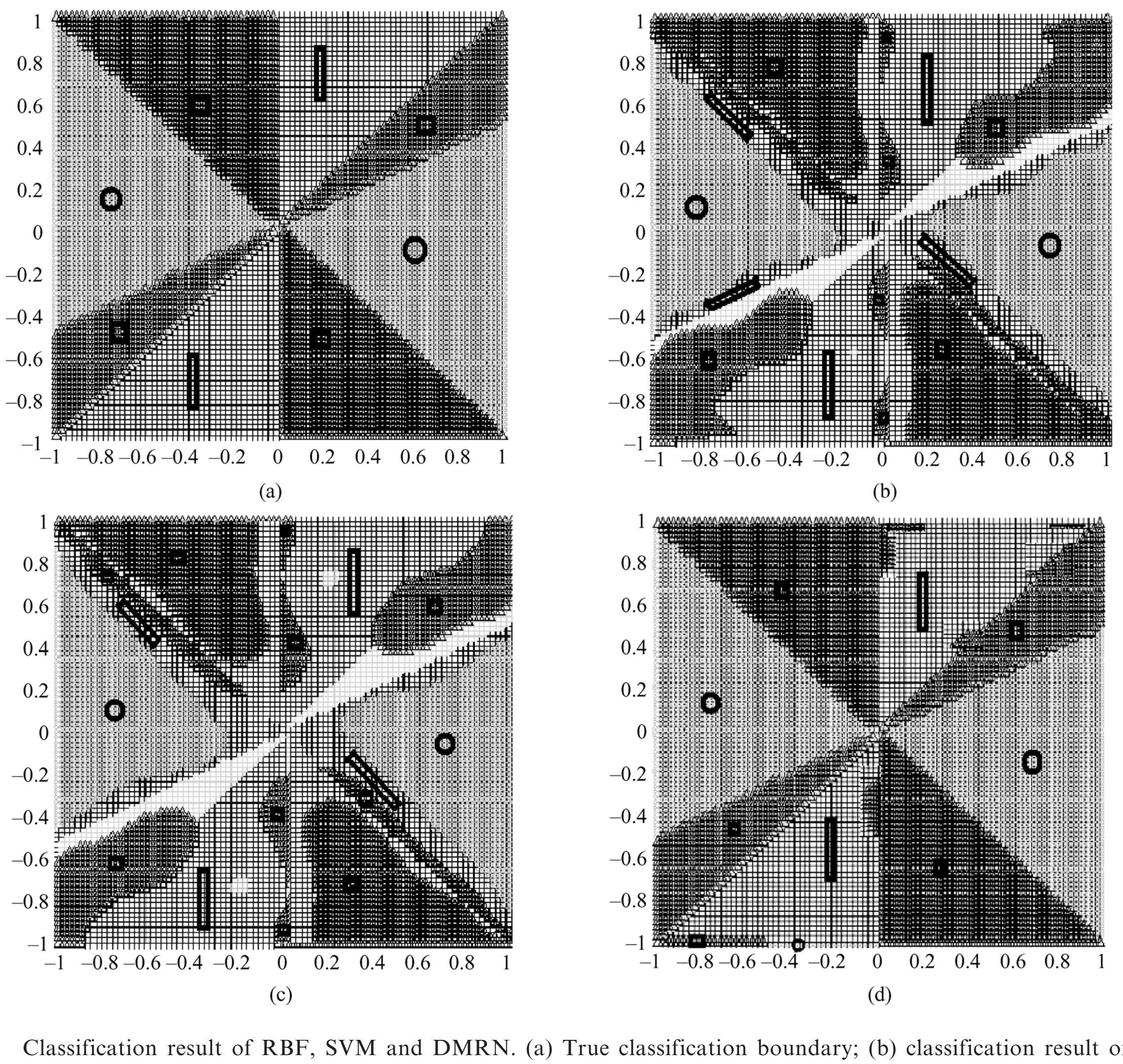

Fig. 2 Classification result of RBF, SVM and DMRN. (a) True classification boundary; (b) classification result of RBF; (c) classification result of SVM; (d) classification result of DMRN

ridgelet network is more suitable for processing high dimensional data by avoiding the complex construction of high dimensional wavelets. Third, the ridgelet frame theory makes a more flexible construction of the network, so we can adjust the structure of the network easily. Finally, it can adopt some ridgelets to approximate an input-output mapping, and avoid the high complexity decomposition in the case of large number of samples. In the paper some theoretical suggestions on its construction are also given, and some theorems are presented to prove its approximation capability and approximation rate, and the experiment results are also presented to prove its efficiency.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 60601029), the National Basic Research Program of China (No. A 50105020025).

References

1. McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 1943, 5: 115–133
2. Park J, Sandberg I W. Universal approximation using radial-basis-function networks. Neural Computation, 1991, 3(2): 246–257
3. Zhang Q, Benveniste A. Wavelet networks. IEEE Transactions on Neural Networks, 1992, 3(6): 889–898
4. Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of Physiology, 1962, 160: 106–154
5. Candès E J. Harmonic analysis of neural networks. Applied and Computational Harmonic Analysis, 1999, 6(2): 197–218
6. Donoho D L. Tight frames of k-plane ridgelets and the problem of representing objects that are smooth away from d-dimensional singularities. In: Proceedings of the National Academy of Sciences, USA, 1999, 96(5): 1828–1833
7. Candès E J, Donoho D L. Ridgelets: a key to higher-dimensional intermittency? Philosophical Transactions of the Royal Society of London Series A, 1999, 357(1760): 2495–2509
8. Starck J L, Candès E J, Donoho D L. The curvelet transform for image denoising. IEEE Transactions on Image Processing, 2002, 11(6): 670–684
9. Grochenig K. Acceleration of the Frame Algorithm. IEEE Transactions on Signal Processing, 1993, 41(12): 3331–3340
10. Candès E J. Ridgelet: theory and applications. Dissertation for the Doctoral Degree. CA: Stanford University, 1998
11. Daubechies I. The wavelet transform: time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 1990, 36(5): 961–1005
12. Donoho D L. Orthonormal ridgelets and linear singularities. SIAM Journal on Mathematical Analysis, 2000, 31(5): 1062–1099
13. Do M N, Vetterli M. The finite ridgelet transform for image representation. IEEE Transactions on Image Processing, 2003, 12(1): 16–28
14. Lin W, Kovvali N, Carin L. Ridgelet-based implementation of multi-resolution time domain. IEEE Transactions on Antennas and Propagation, 2005, 53(8): 2688–2699
15. Hou B, Liu F, Jiao L C. Linear feature detection based on ridgelet. Science in China (Series E), 2003, 46(2): 141–152
16. Yang S Y, Jiao L C, Wang M. An adaptive ridgelet neural network model. Journal of Xidian University, 2005, 32(6): 890–894 (in Chinese)