

# 基于小波框架的自适应径向基函数网络<sup>1)</sup>

张茁生 刘贵忠 张宗平

(西安交通大学电信学院信息与通信工程系 西安 710049)

(E-mail: zszhang@xjtu.edu.cn)

**摘要** 给出了由高斯径向基函数生成的一组小波框架, 建立在小波框架理论的基础上, 构造性地证明了高斯径向基函数网络可以任意精度地逼近  $L^2(R^d)$  中的函数. 在此基础上, 利用高斯径向基函数的时频局部化性质和自适应投影原理, 进一步给出了构造和训练网络的自适应学习算法. 应用到信号的重构和去噪, 获得了良好的效果.

**关键词** 径向基函数网络, 小波框架, 普遍近似性, 自适应学习算法

**中图分类号** TP273.2

## RADIAL BASIS FUNCTION NETWORKS BASED ON WAVELET FRAMES

ZHANG Zhuo-Sheng LIU Gui-Zhong ZHANG Zong-Ping

(Department of Information & Communication Engineering, School of Electrical and Information Engineering,  
Xi'an Jiaotong University, Xi'an 710049)

(E-mail: zszhang@xjtu.edu.cn)

**Abstract** A set of wavelet frames generated by Gaussian radial basis functions are presented. It is constructively proved that a radial basis function network with Gaussian activation functions can approximate any function in  $L^2(R^d)$  with desired accuracy. Furthermore, an adaptive learning algorithm for constructing and training networks is proposed based on time-frequency localization properties of Gaussian radial basis functions and the adaptive projection algorithm. Applications to signal reconstruction and noise elimination are given.

**Key words** Radial function networks, wavelet frames,  $L^2(R^d)$  approximation property, adaptive learning algorithms

## 1 引言

径向基函数(Radial Basis Function, RBF)网络以其与众不同的性质、简单的网络结构和有效的训练过程已越来越广泛地受到人们的欢迎. 已证明, 具有足够多隐层结点的RBF网络可以任意精度地逼近  $L^2(R^d)$  中的函数<sup>[1~3]</sup>. 然而, 我们注意到, 这些已有的证明

1) 国家自然科学基金(69872030)、教育部优秀青年教师基金(97年度)资助

收稿日期 1999-12-22 收修改稿日期 2000-12-01

是从理论上给出的,基本上是非构造性的,即这些证明方法不能自然地给出网络的系统构造.关于 RBF 网络的学习算法,大多数分为两个步骤进行<sup>[4~7]</sup>.其一是确定隐层结点的中心和宽度,它是在隐层结点个数预先给定的情况下,利用有监督和无监督方式进行的;其二是学习连接隐层与输出层的权值,它是一个简单的二次凸函数的优化问题.但是,在这些学习算法中,网络隐层结点的数目是预先给定的,这常常导致网络的体积过大,进而增加网络的复杂性和消耗大量的计算时间,影响网络的泛化能力.为此,人们已给出一些实际的方法确定网络隐层结点的数目和中心的位置,主要是采用构造和删除的方法,或是开始使用一个隐层结点,然后根据需要增加新的结点,或是开始使用很多隐层结点,然后从中删除冗余的隐层结点<sup>[8~12]</sup>,其中最值得提出的是用于构造和训练 RBF 网络的正交最小二乘学习算法<sup>[10]</sup>.径向基函数网络是一个单隐层的前馈型神经网络,其结构如图 1 所示.

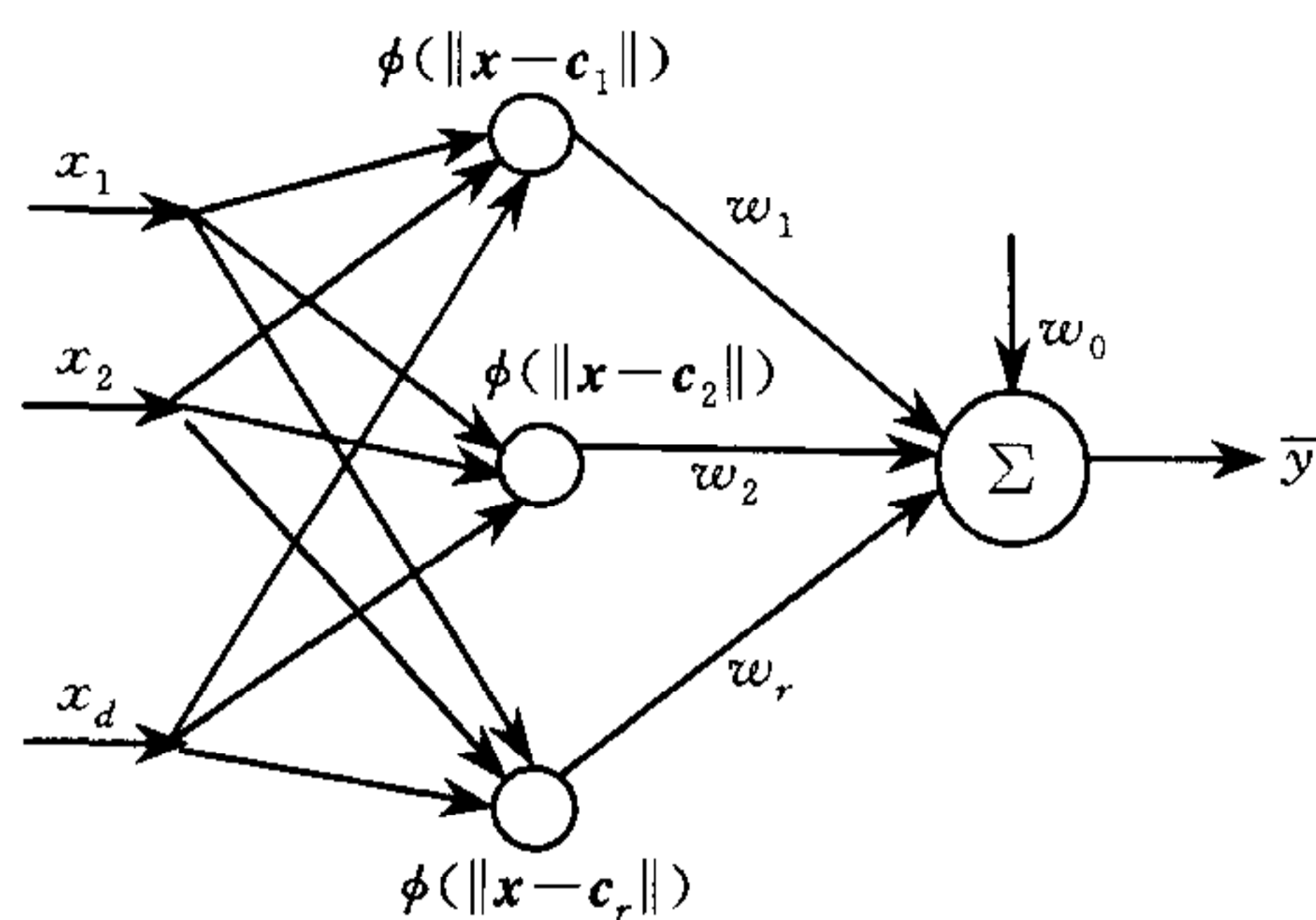


图 1 径向基函数网络的结构

这个网络的输出为

$$\bar{y} = \omega_0 + \sum_{i=1}^r \omega_i \phi(\|x - c_i\|),$$

式中  $x = (x_1, x_2, \dots, x_d) \in R^d$  是网络的输入向量,  $\|\cdot\|$  是  $R^d$  中的欧几里得范数,  $\phi(\cdot): R^+ \rightarrow R$  为径向基函数,  $\omega_i, 1 \leq i \leq r$ , 是网络的权系数,  $c_i \in R^d, 1 \leq i \leq r$ , 是径向基函数的中心,  $r$  是径向基函数的数目.

本文从一个新的观点研究 RBF 网络的逼近能力、构造和学习算法.第二部分介绍了

小波框架的基本概念,给出了小波函数生成框架的一个充分条件;第三部分给出了由高斯函数生成的一组小波框架和高斯径向基函数网络可以任意逼近  $L^2(R^d)$  中函数的构造性证明;第四部分给出了构造和训练网络的自适应学习算法;第五部分给出了在信号的重构和去噪等方面的应用;第六部分是结论.

## 2 小波框架

**定义 1**<sup>[13]</sup>. 设  $H$  为 Hilbert 空间,  $\{h_n\} \subset H$ , 如果存在常数  $A > 0, B < \infty$ , 使

$$A \|f\|^2 \leq \sum_n |\langle f, h_n \rangle|^2 \leq B \|f\|^2, \quad \forall f \in H \quad (1)$$

称  $\{h_n\}$  为  $H$  中的框架,  $A$  和  $B$  分别为框架的下界和上界.

在通常情况下,框架为一组“冗余的基”. 易知<sup>[14]</sup>,  $\{h_n\}$  在  $H$  中稠, 即  $\overline{\text{span}}\{h_n\} = H$ .

设  $\psi \in L^2(R^d)$  为小波函数, 令

$$\psi_{mn}(x) = a^{-\frac{1}{2}dm} \psi(a^{-m}x - nb), \quad m \in Z, n \in Z^d \quad (2)$$

其中  $a > 1, b > 0$ , 如果  $\{\psi_{mn}\}$  构成  $L^2(R^d)$  中的框架, 则称  $\{\psi_{mn}\}$  为小波框架<sup>[15]</sup>.

仿 Daubechies 在文献[14]中的结果, 根据文献[15]中定理 1 的证明, 立即有下面定理. 它可使用数值的方法确定平移参数  $b_0$ , 使当  $0 < b < b_0$  时,  $\{\psi_{mn}\}$  构成小波框架, 并且给出了框架界  $A$  和  $B$  的估计.

**定理 1.** 设  $a > 1$ ,  $\psi \in L^2(R^d)$  为小波函数, 且满足

$$a) m(\psi, a) = \operatorname{ess\,inf}_{\|\omega\| \in (1, a]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a^m \omega)|^2 > 0;$$

b) 存在  $\epsilon > 0$ , 使

$$\sup_{s \in \mathbb{Z}^d} [(1 + s^T s)^{d(1+\epsilon)/2} \beta(s)] < \infty,$$

其中  $\beta(s) = \sup_{\|\omega\| \in (1, a]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a^m \omega)| |\hat{\psi}(a^m \omega + s)|$ . 取

$$b_0 = \sup \left\{ b \left| m(\psi, a) - \sum_{\substack{k \in \mathbb{Z}^d \\ k \neq 0}} \left[ \beta\left(\frac{2\pi}{b} k\right) \beta\left(-\frac{2\pi}{b} k\right) \right]^{1/2} > 0 \right. \right\} \quad (3)$$

则当  $0 < b < b_0$  时,  $\{\psi_{mn}\}$  构成  $L^2(R^d)$  中的框架, 并且有如下框架界估计

$$A = \left(\frac{2\pi}{b}\right)^d \left\{ m(\psi, a) - \sum_{\substack{k \in \mathbb{Z}^d \\ k \neq 0}} \left[ \beta\left(\frac{2\pi}{b} k\right) \beta\left(-\frac{2\pi}{b} k\right) \right]^{1/2} \right\} \quad (4)$$

$$B = \left(\frac{2\pi}{b}\right)^d \left\{ M(\psi, a) + \sum_{\substack{k \in \mathbb{Z}^d \\ k \neq 0}} \left[ \beta\left(\frac{2\pi}{b} k\right) \beta\left(-\frac{2\pi}{b} k\right) \right]^{1/2} \right\} \quad (5)$$

其中  $M(\psi, a) = \operatorname{ess\,sup}_{\|\omega\| \in (1, a]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a^m \omega)|^2$ .

需要指出的是, 文献[15]中定理 1 的条件(b)是多余的. 因为在条件(c)中令  $s = \mathbf{0}$ , 立即得到条件(b).

### 3 高斯函数生成的小波框架、RBF 网络对 $L^2(R^d)$ 中函数的逼近性

设  $\phi(x) = e^{-\|x\|^2/2\sigma^2}$ ,  $x \in R^d$ ,  $\sigma > 0$ , 为高斯径向基函数, 令

$$\psi(x) = 1/\sigma_1 e^{-\|x\|^2/2\sigma_1^2} - 1/\sigma_2 e^{-\|x\|^2/2\sigma_2^2}, \quad x \in R^d, \sigma_1, \sigma_2 > 0, \sigma_1 \neq \sigma_2 \quad (6)$$

则

$$\hat{\psi}(\omega) = (2\pi)^{-n/2} \int_{R^d} \psi(x) e^{-j\omega^T x} dx = e^{-\sigma_1^2 \|\omega\|^2/2} - e^{-\sigma_2^2 \|\omega\|^2/2} \quad (7)$$

显然,  $\hat{\psi}(\mathbf{0}) = 0$ , 且  $\psi$  和  $\hat{\psi}$  均为指数衰减的, 故小波函数  $\psi$  具有极佳时频局部化性质. 特别, 当取  $\sigma_1 = 1, \sigma_2 = 2$  时, 有

$$\psi(x) = e^{-\|x\|^2/2} - 1/2 e^{-\|x\|^2/8} \quad (8)$$

$$\hat{\psi}(\omega) = e^{-\|\omega\|^2/2} - e^{-2\|\omega\|^2} \quad (9)$$

图 2 给出了  $d=1$  和  $d=2$  时  $\psi$  和  $\hat{\psi}$  的图形.

下面证明, 当适当地选取  $b$  时,  $\{\psi_{mn}\}$  构成  $L^2(R^d)$  中的框架.

**定理 2.** 设  $a > 1$ ,  $b_0$  和  $\psi$  分别由式(3)和式(6)定义, 则  $b_0 > 0$ , 且当  $0 < b < b_0$  时,  $\{\psi_{mn}\}$  构成  $L^2(R^d)$  中的框架.

**证明.** 不妨设  $\sigma_2 > \sigma_1$ . 任取  $\omega \in R^d$ ,  $\|\omega\| \in (1, a]$ , 有  $\sum_{m \in \mathbb{Z}} |\hat{\psi}(a^m \omega)|^2 \geq |\hat{\psi}(\omega)|^2 \geq e^{-\sigma_1^2 a^2/2} (1 - e^{-(\sigma_2^2 - \sigma_1^2)/2})$ , 故  $m(\psi, a) > 0$ . 其次证明, 存在常数  $C$ , 使

$$|\hat{\psi}(\omega)| \leq C \|\omega\|^2 (1 + \|\omega\|^2)^{-(d+3)/2}, \quad \omega \in R^d \quad (10)$$

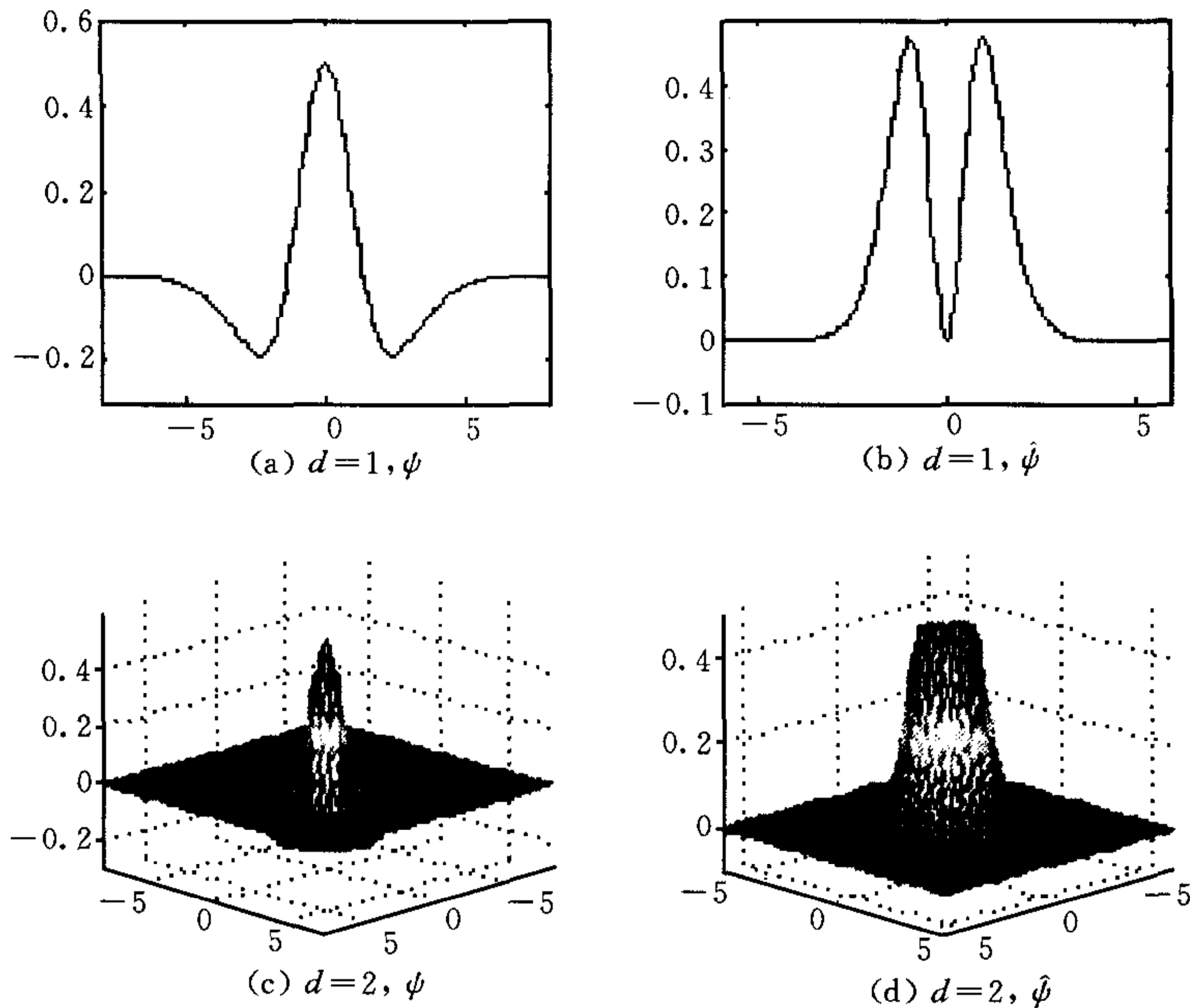


图 2  $d=1$ 和 $d=2$ 时 $\psi$ 和 $\hat{\psi}$ 的图形

由于  $\lim_{\|\omega\| \rightarrow 0} \hat{\psi}(\omega) / \|\omega\|^2 = (\sigma_2^2 - \sigma_1^2) / 2$ , 故存在常数  $C_1 > 0, M_1 > 0$ , 使当  $\|\omega\| < M_1$  时, 有  $|\hat{\psi}(\omega)| \leq C_1 \|\omega\|^2$ , 于是  $|\hat{\psi}(\omega)| \leq C_1 (1 + M_1^2)^{(d+3)/2} \|\omega\|^2 (1 + \|\omega\|^2)^{-(d+3)/2}, \omega \in R^d, \|\omega\| < M_1$ . 另一方面, 由于  $\lim_{\|\omega\| \rightarrow \infty} \hat{\psi}(\omega) / (1 + \|\omega\|^2)^{-(d+3)/2} = 0$ , 存在常数  $C_2, M_2 > 0, M_2 > M_1$ , 当  $\|\omega\| > M_2$  时, 有  $|\hat{\psi}(\omega)| \leq C_2 (1 + \|\omega\|^2)^{-(d+3)/2}$ . 从而  $|\hat{\psi}(\omega)| \leq C_2 / M_1^2 \|\omega\|^2 (1 + \|\omega\|^2)^{-(d+3)/2}, \omega \in R^d, \|\omega\| > M_2$ . 由于  $\hat{\psi}(\omega)$  在  $[M_1, M_2]$  上连续, 存在常数  $C_3 > 0$ , 使  $|\hat{\psi}(\omega)| \leq C_3, \|\omega\| \in [M_1, M_2]$ . 因此  $|\hat{\psi}(\omega)| \leq (C_3 (1 + M_2^2)^{(d+3)/2} / M_1) \|\omega\|^2 (1 + \|\omega\|^2)^{-(d+3)/2}, \omega \in R^d, \|\omega\| \in [M_1, M_2]$ . 综合上述, 式(10)成立.

最后证明 
$$\beta(s) = \sup_{\|\omega\| \in (1, a]} \sum_{m \in Z} |\hat{\psi}(a^m \omega)| |\hat{\psi}(a^m \omega + s)| \leq \tilde{C} (1 + \|s\|^2)^{-(d+1)/2}, \tilde{C} > 0, s \in R^d \tag{11}$$

任取  $s, \omega \in R^d, \|\omega\| \in (1, a]$ , 据式(10), 有  $\sum_{m \in Z} |\hat{\psi}(a^m \omega)| |\hat{\psi}(a^m \omega + s)| \leq C^2 \sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|a^m \omega + s\|^2)^{-(d+1)/2} + C^2 \sum_{m=0}^{\infty} [(1 + \|a^m \omega\|^2)(1 + \|a^m \omega + s\|^2)]^{-(d+1)/2}$ . 当  $\|s\| \geq 2$  时, 有  $\|a^m \omega + s\| \geq \|s\| - \|a^m \omega\| \geq \|s\| - 1 \geq \|s\| / 2$ , 从而  $\sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|a^m \omega + s\|^2)^{-(d+1)/2} \leq 4^{-(d+1)/2} \sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|s\|^2)^{-(d+1)/2}$ ; 当  $\|s\| < 2$  时, 有  $1 + \|a^m \omega + s\|^2 \geq 1 \geq (1 + \|s\|^2) / 5$ , 于是  $\sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|a^m \omega + s\|^2)^{-(d+1)/2} \leq 5^{(d+1)/2} \sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|s\|^2)^{-(d+1)/2}$ . 综合上述, 存在常数  $\tilde{C}_1 > 0$ , 使  $\sum_{m=-\infty}^{-1} a^{2(m+1)} (1 + \|a^m \omega + s\|^2)^{-(d+1)/2} \leq \tilde{C}_1 (1 + \|s\|^2)^{-(d+1)/2}$ . 注意到  $(1 + \|x - y\|^2)(1 + \|x + y\|^2) \geq 1 + \|y\|^2, \forall x, y \in R^d$ , 从而

$(1 + \|a^m \omega\|^2)(1 + \|a^m \omega + s\|^2) \geq 1 + \|s\|^2/4 \geq (1 + \|s\|^2)/4$ . 因此,有 $[(1 + \|a^m \omega\|^2)(1 + \|a^m \omega + s\|^2)]^{-(d+1)/2} \leq [(1 + \|a^m \omega\|^2)(1 + \|a^m \omega + s\|^2)]^{-1/4} 4^{(d+1/2)/2} (1 + \|s\|^2)^{-(d+1/2)/2} \leq 4^{(d+1/2)/2} a^{-m/2} (1 + \|s\|^2)^{-(d+1/2)/2}$ . 从而,存在常数  $\tilde{C}_2 > 0$ , 使  $\sum_{m=0}^{\infty} [(1 + \|a^m \omega\|^2)(1 + \|a^m \omega + s\|^2)]^{-(d+1)/2} \leq \tilde{C}_2 (1 + \|s\|^2)^{-(d+1/2)/2}$ . 综合上述,式(11)得证. 进而,由定理 1 知,定理得证. 证毕.

对于式(8)确定的小波函数  $\psi$ ,取  $a=2.0$ , 据式(3)得:在一维情况下,当  $0 < b < 3.8$  时,  $\{\psi_{mn}\}$  构成  $L^2(R)$  中的小波框架;在二维情况下,当  $0 < b < 1.8$  时,  $\{\psi_{mn}\}$  构成  $L^2(R^2)$  中的小波框架. 表 1 和表 2 分别给出了一维和二维情况下,伸缩步长  $b$  的一些值所对应的框架界  $A, B$  和  $B/A$  的值.

表 1

$b$	$A$	$B$	$B/A$
0.5	3.9850	4.1059	1.0303
1.0	1.9925	2.0529	1.0303
1.5	1.3244	1.3725	1.0363
2.0	0.9312	1.0951	1.1721
2.5	0.5836	1.0346	1.7730
3.0	0.2954	1.0531	3.5652
3.5	0.0917	1.0641	11.6019

表 2

$b$	$A$	$B$	$B/A$
0.25	200.3088	206.3834	1.0303
0.50	50.0772	51.5958	1.0303
0.75	22.2565	22.9315	1.0303
1.00	12.5105	12.9078	1.0318
1.25	7.7910	8.4767	1.0880
1.50	4.3844	6.9126	1.5766
1.75	1.0737	7.2261	6.7301

**定理 3.** 高斯径向基函数网络可以任意精度地逼近  $L^2(R^d)$  中的任意函数  $f$ . 更确切地说,任取  $\epsilon > 0$ , 存在  $M_0, M_1 \in Z, N_0, N_1 \in Z^d$ , 使

$$\left\| f - \sum_{(m,n) \in Q} \langle f, T^{-1} \psi_{mn} \rangle \psi_{mn} \right\| < \epsilon,$$

式中  $Q = \{(m, n) \mid M_0 \leq m \leq M_1, N_0 \leq n \leq N_1, m \in Z, n \in Z^d\}$ . 我们约定,  $\forall x \in (x_1, x_2, \dots, x_d), y = (y_1, y_2, \dots, y_d) \in Z^d, x \leq y \Leftrightarrow x_i \leq y_i, 1 \leq i \leq d$ .

**证明.** 任取  $g \in L^2(R^d)$ , 令  $Tg = \sum_{m \in Z, n \in Z^d} \langle g, \psi_{mn} \rangle \psi_{mn}$ . 由于  $\{\psi_{mn}\}$  构成  $L^2(R^d)$  中的小波框架,熟知<sup>[13]</sup>,  $T: L^2(R^d) \rightarrow L^2(R^d)$  为有界可逆线性算子,并且  $\{T^{-1} \psi_{mn}\}$  亦构成  $L^2(R^d)$  中的小波框架. 因此,存在常数  $0 < A \leq B < \infty$ , 使

$$A \|g\|^2 \leq \sum_{m \in Z, n \in Z^d} |\langle g, \psi_{mn} \rangle|^2 \leq B \|g\|^2, \quad \forall g \in L^2(R^d) \tag{12}$$

$$B^{-1} \|g\|^2 \leq \sum_{m \in Z, n \in Z^d} |\langle g, T^{-1} \psi_{mn} \rangle|^2 \leq A^{-1} \|g\|^2, \quad \forall g \in L^2(R^d) \tag{13}$$

又由式(13)知,存在  $M_0, M_1 \in Z, N_0, N_1 \in Z^d$ , 使  $\sum_{(m,n) \in (Z, Z^d) \setminus Q} |\langle f, T^{-1} \psi_{mn} \rangle|^2 < \epsilon^2/B$ . 于是

$$\begin{aligned} \left\| f - \sum_{(m,n) \in Q} \langle f, T^{-1} \psi_{mn} \rangle \psi_{mn} \right\| &= \sup_{g \in L^2(R^d), \|g\|=1} \left| \left\langle \sum_{(m,n) \in (Z, Z^d) \setminus Q} \langle f, T^{-1} \psi_{mn} \rangle \psi_{mn}, g \right\rangle \right| = \\ &= \sup_{g \in L^2(R^d), \|g\|=1} \left| \sum_{(m,n) \in (Z, Z^d) \setminus Q} \langle f, T^{-1} \psi_{mn} \rangle \langle \psi_{mn}, g \rangle \right| \leq \\ &= \left( \sum_{(m,n) \in (Z, Z^d) \setminus Q} |\langle f, T^{-1} \psi_{mn} \rangle|^2 \right)^{1/2} \sup_{g \in L^2(R^d), \|g\|=1} \left( \sum_{(m,n) \in (Z, Z^d) \setminus Q} |\langle \psi_{mn}, g \rangle|^2 \right)^{1/2} < \epsilon, \end{aligned}$$

故定理得证.

证毕.

综合上述,我们构造性地证明了高斯径向基函数网络可以任意精度地逼近  $L^2(R^d)$  中的任意函数.

#### 4 构造和训练径向基函数网络的自适应学习算法

下面讨论在实际应用中,怎样具体地构造和训练网络.由于考虑的小波函数  $\psi$  是两个不同宽度高斯径向基函数的差,高斯径向基函数网络的设计问题就完全等价于以  $\psi$  为隐层结点激活函数的网络的设计问题.

设  $\{(x_i, y_i)\}_{i=1}^N$  为一组训练样本,并且满足  $y_i = f(x_i) + n_i$ ,  $i = 1, 2, \dots, N$ ,  $f \in L^2(R^d)$  为要估计的信号,  $x_i \in R^d$ ,  $y_i \in R$ ,  $n_i \sim N(0, \sigma^2)$  为高斯白噪声序列,则构造和训练网络的自适应学习算法可作如下描述.

##### 1) 对小波框架进行截取

把由无限多个元组成的小波框架  $\{\psi_{mn}\}$  截成有限集.截取基于下面两个方面的考虑:a) 由于训练样本  $\{(x_i, y_i)\}_{i=1}^N$  是有限的,  $f$  仅可被估计到某个尺度细节.再注意到  $\hat{\psi}$  是指数衰减的,故更精细的尺度应当被截取,从而得到一个尺度的取值范围 ( $m_1 \leq m \leq m_2$ ); b) 由于  $\{x_i\}_{i=1}^N$  在输入空间  $R^d$  的一个有限域  $D$  内和  $\psi$  的指数衰减性,支撑域不包含  $\{x_i\}_{i=1}^N$  的小波应当被截去.具体作法为

令

$$G = \{x \in R^d \mid |\psi(x)| > \varepsilon \max_{x \in R^d} |\psi(x)|\}, \quad T_\varepsilon = \sup_{x \in G} \|x\|,$$

其中  $\varepsilon$  是一个很小的正数.定义小波函数  $\psi$  的支撑域为

$$S_{00} = \{x \in R^d \mid \|x\| < T_\varepsilon\},$$

于是,  $\psi_{mn}$  的支撑域为

$$S_{mn} = \{x \in R^d \mid \|x - 2^m nb\| < 2^m T_\varepsilon\}.$$

令

$$I = \{(m, n) \mid S_{mn} \cap D \neq \phi, m_1 \leq m \leq m_2\},$$

选取网络的隐层结点的神经元分别为所有  $\{\psi_{mn} \mid (m, n) \in I\}$  中的小波函数,从而完成网络的初步设计.

##### 2) 用自适应投影算法化简网络

在上面选取的小波基元的过程中,相当于在时频平面上对小波框架  $\{\psi_{mn}\}$  进行正规的截取.如果用截取后的框架元去逼近信号  $f$ ,那么精细尺度的小波(高频小波)可能在信号的整个输入区域中被使用.再注意到小波函数的时频局部化特性,这就常常导致被选用的小波基元是非常冗余的.为此,使用自适应投影算法<sup>[16]</sup>进一步化简网络.

为了叙述简单起见,把  $\{\psi_{mn} \mid (m, n) \in I\}$  重记为  $\{\psi_1, \psi_2, \dots, \psi_L\}$ . 为从  $\{\psi_1, \psi_2, \dots, \psi_L\}$  中选取子集“最佳”地逼近信号  $f$ ,把  $\psi_i$ ,  $1 \leq i \leq L$  视为回归因子.令

$$Y = (y_1, y_2, \dots, y_N)^T, \quad \Psi_i = (\psi_i(x_1), \psi_i(x_2), \dots, \psi_i(x_N)),$$

于是,问题就转变为从  $\{\Psi_1, \Psi_2, \dots, \Psi_L\}$  中选取子集使它“最佳”地逼近  $Y$ .为此,对  $Y$  在基集  $\{\Psi_1, \Psi_2, \dots, \Psi_L\}$  上实施自适应投影算法.具体算法描述如下.

Step 0. 把  $\{\Psi_i\}_{i=1}^L$  进行能量归一化:  $\Psi'_i = \Psi_i / \|\Psi_i\|, i=1, 2, \dots, L.$

Step 1. 求  $\Psi'_{i_0} \in \{\Psi'_i\}_{i=1}^L$  使  $|\langle Y, \Psi'_{i_0} \rangle| = \max_{1 \leq i \leq L} |\langle Y, \Psi'_i \rangle|$ ;  $Y$  分解为  $Y = \langle Y, \Psi'_{i_0} \rangle \Psi'_{i_0} + R_Y^1.$

Step 2. 求  $\Psi'_{i_k} \in \{\Psi'_i\}_{i=1}^L$ , 使  $|\langle R_Y^k, \Psi'_{i_k} \rangle| = \max_{1 \leq i \leq L} |\langle R_Y^k, \Psi'_i \rangle|$ ;  $R_Y^k$  分解为  $R_Y^k = \langle R_Y^k, \Psi'_{i_k} \rangle \Psi'_{i_k} + R_Y^{k+1}.$

Step 3. 如果残差  $\|R_Y^{k+1}\| / \|Y\| < \epsilon$  或残差衰减率  $\|R_Y^k - R_Y^{k+1}\| / \|R_Y^k\| = |\langle R_Y^k, \Psi'_{i_k} \rangle| / \|R_Y^k\| < \epsilon$  ( $\epsilon$  为一较小的正数), 停; 否则, 令  $k = k + 1$ , 转 Step 2.

设迭代选取的基向量为  $\Psi'_{i_0}, \Psi'_{i_1}, \dots, \Psi'_{i_{r-1}}$ , 则选取小波网络隐层结点的神经元函数为  $\psi_{i_0}, \psi_{i_1}, \dots, \psi_{i_{r-1}}$ . 这时所对应的网络权系数分别为  $\langle R_Y^0, \Psi'_{i_0} \rangle / \|\Psi_{i_0}\|, \langle R_Y^1, \Psi'_{i_1} \rangle / \|\Psi_{i_1}\|, \dots, \langle R_Y^{r-1}, \Psi'_{i_{r-1}} \rangle / \|\Psi_{i_{r-1}}\|.$

### 5 数值实验

在实验中,使用小波函数

$$\psi(x) = e^{-\|x\|^2/2} - 1/2e^{-\|x\|^2/8}, \quad x \in R^d,$$

取  $a=2.0, b=0.5.$

#### 1) 信号重构

用 Bumps 信号和 Doppler 信号作为测试信号. 记原始信号为  $f(x), 0 \leq x \leq 1.$  设采样序列为  $y_i = f(x_i), i=1, 2, \dots, N, N=1024,$  其中  $\{x_i\}_{i=1}^N$  为  $[0, 1]$  上的均匀采样, 则通过自适应投影迭代法迭代  $n$  次后的重构信号为

$$\tilde{f}(x) = \sum_{k=0}^{n-1} (\langle R_Y^k, \Psi_{i_k} \rangle / \|\Psi_{i_k}\|) \psi_{i_k}(t),$$

取自适应投影迭代法的迭代终止准则为残差  $\|R_Y^{k+1}\| / \|Y\| < \epsilon, \epsilon=0.025,$  我们构造了分别具有 81 个和 93 个隐层结点的以小波函数为激活函数的网络, 即具有 181 个和 186 个隐层结点的高斯径向基函数网络, 图 3 给出了两个测试信号的原始信号和重构信号.

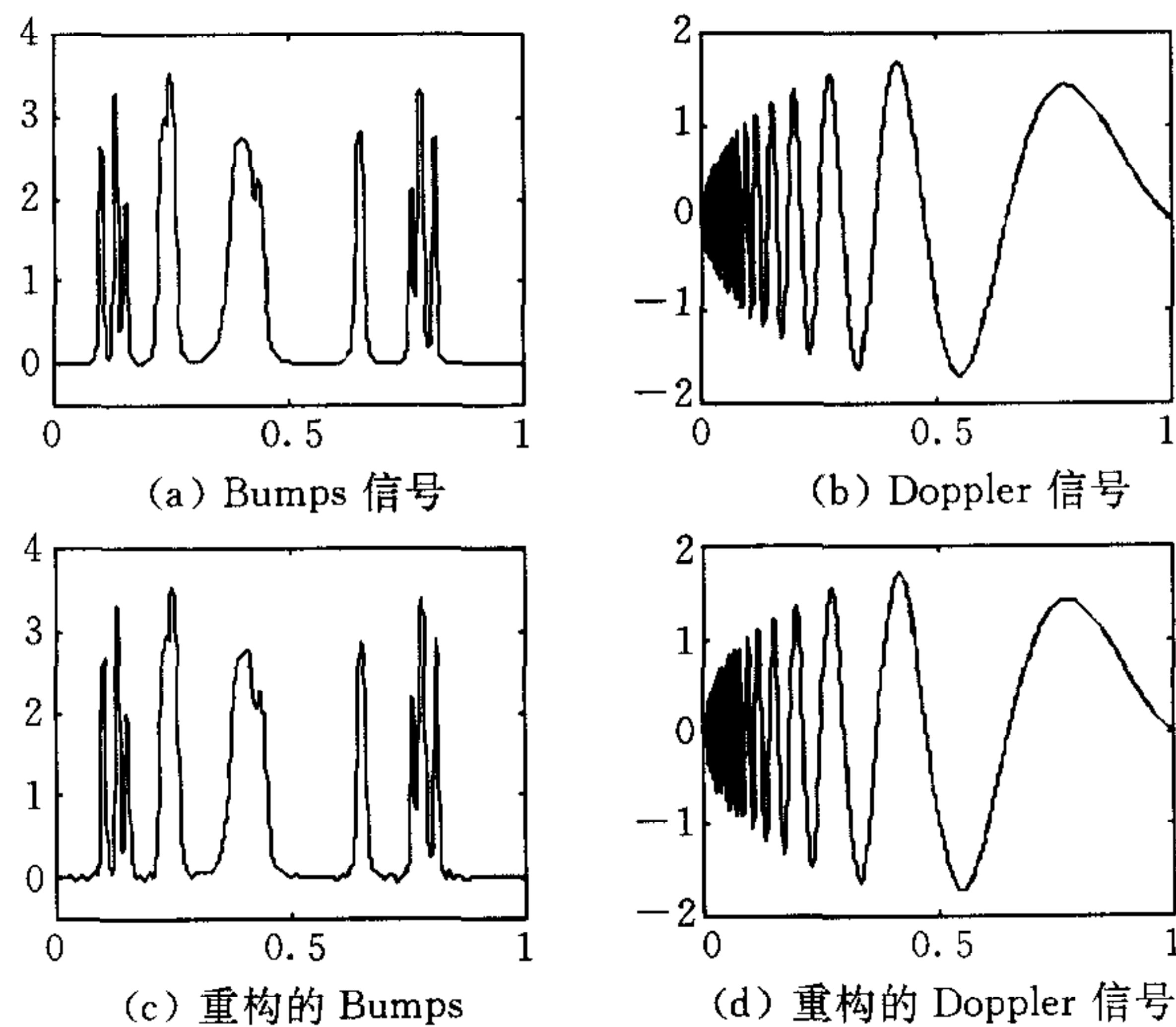


图 3 两个测试信号的原始信号和重构信号

## 2) 信号去噪

本文构造的高斯径向基函数网络具有非常好的抗噪性能. 这是因为在时频平面上, 信号的能量是集中的, 而随机噪声的能量是均匀分布的, 以及高斯径向基函数具有极佳的时频汇聚性能, 故自适应投影算法只有将能量集中的信号全部提取完后, 才会轮到搜索时频平面上能量均匀分布的噪声. 因此, 可利用高斯径向基函数网络进行消噪.

仍然用 Bumps 信号和 Doppler 信号作为测试信号. 设含噪的样本序列为  $y_i = f(x_i) + n_i$ ,  $t = 1, 2, \dots, N$ ,  $N = 1024$ ,  $n_i \sim N(0, \sigma^2)$  为高斯被白噪声, 定义输入信噪比和输出信噪比分别为

$$SNR_{input} = 20 \log_{10} (std(\{f(x_i)\}) / std(\{n_i\})),$$

$$SNR_{output} = 20 \log_{10} (std(\{f(x_i)\}) / std(\{\tilde{f}(x_i) - f(x_i)\})),$$

式中  $std(\cdot)$  表示序列的标准方差,  $\tilde{f}(x)$  为重构信号. 图 4 给出了两个测试信号的去噪处理结果. 取迭代终止准则为残差衰减率  $< \epsilon$ ,  $\epsilon = 0.1$ , 考虑到样本的随机性, 对每一个给定的输入信噪比, 输出信噪比为 100 次不同的输出信噪比的均值. 从图可以看出, 信噪比有较好的改善.

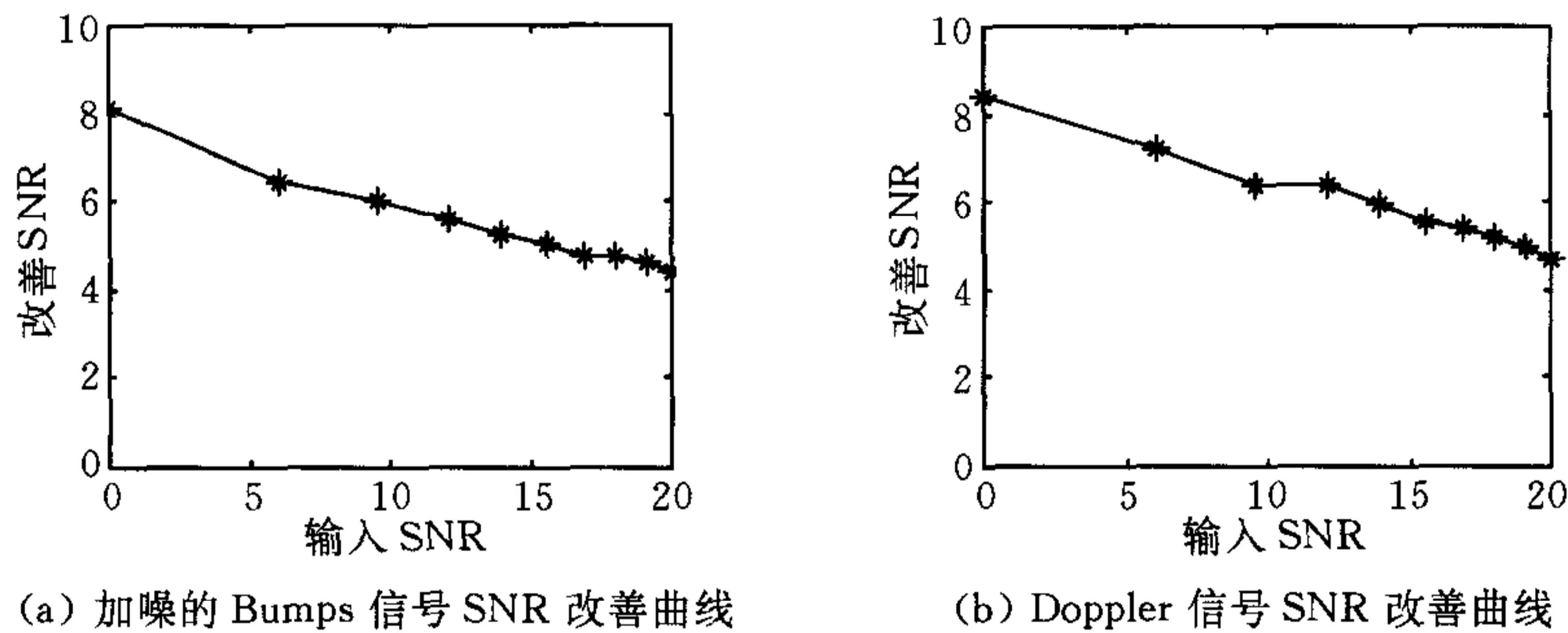


图 4 两个测试信号的去噪处理结果

## 6 结论

本文考虑了高斯径向基函数网络的普遍近似性, 以及构造和训练网络的自适应学习算法. 利用小波框架的理论, 构造出了一组由高斯径向基函数生成的小波框架, 进而构造性地证明了高斯径向基函数网络的普遍近似性. 在此基础上, 进一步给出了一个构造和训练网络的自适应学习算法. 该算法建立在小波框架的理论基础上, 利用高斯小波函数的时频局部化性质, 首先给出了网络的初步设计, 然后用自适应投影学习算法对网络进行化简. 这种学习算法通过迭代自适应地确定网络的隐层结点的个数、RBF 中心的位置和宽度, 以及网络的权系数. 最后, 通过应用到信号重构和去噪以说明该算法的实际效果.

## 参 考 文 献

- 1 Giros F, Poggio T. Networks and the best approximation property. *Biological Cybernetics*, 1990, **63**(1):167~179
- 2 Park J, Sandberg I W. Universal approximation using radial basis function networks. *Neural Computa.*, 1991, **3**(2):246~257
- 3 Park J, Sandberg I W. Approximation and radial basis function networks. *Neural Computa.*, 1993, **5**(3):305~316



- 4 Moody J E, Darken C J. Fast learning in networks of locally-tuned processing units. *Neural Computa.*, 1989, 1(2):281~294
- 5 Karayiannis N B. Gradient descent learning of radial basis neural networks. In: Proc. Int. Conf. Neural Networks, Houston: TX, 1997. 1815~1820
- 6 Karayiannis N B, Mi G W. Growing radial basis neural networks; Merging supervised and unsupervised learning with network growth techniques. *IEEE Trans. Neural Networks*, 1997, 8(6):1492~1506
- 7 Chen S, Gibson G J, Cowan C F N, Grant P M. Reconstruction of binary signals using an adaptive radial-basis-function equalizer. *Signal Processing*, 1991, 22(1):77~93
- 8 Lee S, Kil R M. A Gaussian potential function network with hierarchically self-organizing learning. *Neural Networks*, 1991, 4(2):207~224
- 9 Holcomb T, Morari M. Local training for radial basis function networks; Towards solving the hidden units problem. In: Proc. ACC, Boston: MA, 1991. 2331~2336
- 10 Chen S, Cowan C F N, Grant P M. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Trans. Neural Networks*, 1991, 2(2):302~309
- 11 Fung C F, Billings S A, Luo W. On-line supervised adaptive training using radial function networks. *Neural Networks*, 1996, 9(9):1597~1617
- 12 Whitehead B A, Choat T D. Evolving space-filling curves to distribute radial basis function over an input space. *IEEE Trans. Neural Networks*, 1994, 5(1):15~23
- 13 Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, 1990, 36(2):961~1005
- 14 Daubechies I. Ten Lectures on Wavelets, CBMS-NSF Regional Series in Applied Mathematics, Philadelphia: SIAM, 1992
- 15 Kugarajah T, Zhang Q. Multidimensional wavelet frames. *IEEE Trans. Neural Networks*, 1995, 6(6):1552~1556
- 16 Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Processing*, 1993, 41(12):3397~3415

**张苗生** 副教授, 2001 年获西安交通大学信息与通信工程专业博士学位. 主要研究领域为神经网络、小波分析、联合时频分析及非平稳信号处理等.

**刘贵忠** 教授, 博士生导师, 1989 年获荷兰 Eindhoven 大学博士学位. 主要研究领域为非平稳信号处理、多媒体技术、通信及测控等.