第1章 CHAPTER 1

引言

分析平稳信号的理想工具是傅里叶变换。对于非平稳信号,傅里叶变换不再是有效的分析工具,因为其无法描述信号的局部频率特征。而小波变换正是分析非平稳信号的有力工具。小波变换是傅里叶变换的新发展,它既保留了傅里叶变换的优点,又弥补了傅里叶变换在信号分析上的一些不足。原则上讲,小波变换适用于以往一切傅里叶变换应用的领域。但小波变换并不是万能的,作为一种数学工具,小波变换(分析)有其特定的应用范围,即面向更能发挥小波分析优势的时间(空间)-频率局域性问题。

小波分析作为一门新的数学学科,包含了丰富的数学内容,并推动了泛函分析和调和分析理论的发展,同时,在诸如图像压缩、信号去噪、自适应滤波、数值分析和物理学等领域得到了广泛的应用,是当前最为活跃的应用研究领域之一,并逐渐形成为一门极具生命力的新学科。

1.1 小波理论的发展

从历史上追溯,小波分析的原始思想形成于 20 世纪初,即用一个函数的伸缩及平移构成 $L^2(R)$函数空间的一组基,此函数称为小波。第一个小波是由 Haar 在 1910 年提出的,他在一篇描述抽象 Hibert 空间特性的论文中给出了一个由盒函数产生的 $L^2(R)$函数空间的一组正交基。1936 年 Littlewood 和 Paley 开发出一种利用八度音阶将频率分组的方法,这是按二进制对频率成分进行分组的傅里叶分析思想的最早起源。1946 年,Gabor 提出加窗傅里叶变换,可以反映信号在任意局部范围的频率特性。

经过数学家、物理学家、地理学家半个世纪的共同努力,Haar 系已发展成为统一的理论框架,使小波分析成为傅里叶分析发展史上一个新的里程碑(本章参考文献 1 对前人这些零散的工作有较详尽的论述)。直到 20 世纪 80 年代中期由一批数学家领导的“French School”小组为小波分析奠定了坚实的数学基础,Daubechies 在本章参考文献 2 中很好地总结了小波分析的发展历史。小波变换

是由法国地质物理学家 Morlet 于 1980 年提出的，随后，他与法国理论物理学家 Grossman 共同提出连续小波变换的几何体系，其基础是平移和伸缩下的不变性[3]，这使得能够将一个信号分解成对空间和尺度的独立贡献，同时又不损失原有信号的信息。1982 年，Strömberg 给出了一个无限支撑的、正交的逐段多项式小波。Battle 和 Lemarie 也分别独立地构造了类似的小波。1985 年，法国数学家 Meyer 给出了一个指数衰减的任意阶可导小波。1988 年 Daubechies 基于离散滤波器迭代方法构造了紧支撑标准正交小波基，即一系列的具有任意选定正则性的、有限支撑的、正交的尺度函数和小波，并将当时所有正交小波的构造统一起来，为以后的构造设定了框架[4]。随后她又发表了长篇综述[5]，对小波理论的发展和推广起到积极的作用，该文成为目前小波理论研究的最重要的文献之一。

1989 年 Mallat 将计算机视觉领域内多尺度分析的思想引入到小波分析中，提出多分辨率分析(multiresolution analysis，MRA)的概念，用多分辨率分析来定义小波，给出了构造正交小波基的一般方法和与 FFT 相对应的快速小波算法——Mallat 算法，并将它用于图像分析和完全重构[6]。它使许多以前分散在各应用领域里研究的小波成果有可能统一在同一个理论框架下。

1992 年，在小波变换的基础上，R. R. Coifman 和 M. V. Wickerhauser 进一步提出了小波包(wavelet packet)的概念[7,8]，并从数学上作了严密的推导。利用推广的二尺度方程，原来的尺度函数和小波可以生成一族包括原小波基在内的“小波包”函数。通过引入 Shannon 熵作为信号处理应用中对不同小波包基函数的选定准则，为信号进行自适应频带划分提供了工具[7,9]。

1992 年，A. Cohen 和 I. Daubechies 提出了“双正交小波”的概念，即对同一信号 $f\in L^2(R)$，其分析小波和综合小波可以是两组不同的函数系[10]。

1993 年，基于 M. Vetterli 和 P. P. Vaidyanathan 各自独立提出的多采样率数字信号处理器、广义镜像滤波器组和共轭正交镜像滤波器组的理论[11,12]，产生了 M 带小波，它有一个尺度函数和 $M-1$ 个小波函数[13,14]。

由于实际信号均是有限支撑的，为了很好地刻画信号，避免出现譬如周期化处理引起的边界效应或失真，1993—1994 年，许多学者从不同的角度提出了小波的思想和构造方法[15,16]。

为了解决正交小波基没有线性相位的问题，有学者提出多小波的理论[17~20]，用向量小波来代替标量小波以满足线性相位的要求。1994 年，T. N. T. Goodman 和 S. L. Lee 首先提出了这个概念，并且用 Hermit 样条构造了第一个多小波[19]。而第一个非样条多小波是由 Geronimo，Hardin 和 Massopust 利用分形理论在 1994 年给出的，记为 GHM 多小波——一个著名的多小波[20]。

在多小波的思想出现的同时，1994 年 Sweldens 提出了用提升方法来构造具有线性相位的小波变换[21,22]，进而给出整数可逆的提升框架[23,24]，使得小波变换向实用前进了一大步。

随着小波理论的深入发展，小波应用和产业的建立也成为一种趋势。各种以小波理论研究和应用为主的公司和研究小组相继成立，例如，以 I. Daubeches 为首的 Bell 实验室小波小组、D. L. Donoho 领导的 Stanford 大学的小波与统计学中心、Gopinath 和

Burrus 领导的 Rice 大学的小波与多滤波器组中心以及 Wickerhauser 等的小波包应用研究中心。而“Wavelet Digest”成为 Internet 网上发行最广的专业期刊之一。从 20 世纪 90 年代初，国内不少院校和科研单位相继成立了小波理论和应用研究的专题小组，已取得了部分突出成果。而且，可以实现小波变换的集成芯片也已经问世。现在，有关小波理论的文献已经遍布当代信息科学的所有学科。所有这些都充分显示了小波分析自身的优势和强大的生命力。

1.2 小波分析的应用

小波变换是近十几年信号处理领域研究的一个热点，许多学者将小波在理论上的研究成果应用到诸如图像压缩、特征提取、信号滤波和数据融合等方面，而且小波变换的应用领域还在不断地发展当中。小波之所以在信号处理领域具有很大的优势，在于小波变换可以获得信号的多分辨率描述，这种描述符合人类观察世界的一般规律，同时，小波变换具有丰富的小波基可以适应具有不同特性的信号。

目前，小波在信号处理和图像处理中的应用得到空前发展，大量的理论及应用成果已经出现。

1. 小波分析在图像压缩编码中的应用

由于小波分析具有时频分析、多分辨率分析等优点，易与人类视觉特性相结合，因此小波变换用于信号与图像压缩是小波分析应用的一个重要方面。它的特点是压缩比高，压缩速度快，压缩后能保持信号与图像的特征不变，且在传递过程中可以抗干扰。基于小波分析的压缩方法很多，比较成功的有小波包最优基法、小波域纹理模型方法、小波变换零树压缩、小波变换向量量化压缩等。

根据小波变换的时频局部化原理，基于小波变换的图像编码与经典的 JPEG(moving picture experts group)方法相比，至少具有以下优点[93,94]：

(1) 在基于 DCT(discrete Cosine transform)的图像变换编码中，人们将图像分成 8×8 像素或 16×16 像素的块来处理，故容易出现方块效应。而小波变换是对整幅图像进行变换，因此在重构图像中可以免除采用分块正交变换编码所固有的方块效应。

(2) 小波变换采用塔式分解结构，与人眼由粗到精、由全貌到细节的观察习惯相一致，这是将 WT(wavelet transform)与 HVS(human visual system)的空间分解特性结合起来以改善图像压缩性能的有利条件。小波变换比 DCT 变换更符合人的视觉特性，通过合理的量化编码产生的人为噪声比同样比特率下 JPEG 方法产生的影响要小得多。

(3) 小波变换是图像的时频表示，具有时间频域定位能力，因此可实现图像中平稳成分与非平稳成分的分离，从而可对其进行高效编码。

(4) 尺度由大到小变化，高频图像的细节逐渐增多，因此可以先给出一幅较粗糙的图像，然后根据需要提供更好的细节。

总而言之，小波变换用于图像压缩时，除具有时频局部化分析方法处理非平稳信号的固有长处外，还体现在它具有易于与 HVS 相结合的潜力上。在图像压缩标准“JPEG-

2000”中，小波已成为一个主要的技术。

2. 小波分析在信号滤波中的应用

在传统的基于傅里叶变换的信号处理方法中，要使信号和噪声的频带重叠部分尽可能地小，这样，在频域就可以通过时不变滤波方法将信号同噪声区分开。而当它们的频谱重叠时，这种方法就无能为力了。基于小波变换的非线性滤波是完全不同的，在这种方法中，谱可以重叠，但是谱的幅度(而不是谱的位置)要尽可能不同。在小波变换域，可通过对小波系数进行切削、缩小幅度等非线性处理，以达到滤除噪声的目的。采用这种方法滤波可在一定程度上避免一般低通滤波时造成的信号突变部分变模糊，然而同传统的低通滤波相似，也会造成小程度上的图像(或信号)细节丢失。在使用这种方法时，还应考虑抑制噪声与保留信号细节之间的折衷问题。

3. 小波分析在图像处理边缘检测中的应用

常规的边缘提取只是在原始图像上(时域)进行的，利用图像边缘点处的灰度阶跃变化进行边缘检测，然后提取图像的边缘。常用的边缘检测方法有差分算子法、广义Hough 变换法、最优曲线拟合法以及模糊边缘检测法等[94]。在实际图像中，对应景物图像的边缘灰度变化有时并不十分明显，另外，图像中也存在噪声。因此，时域方法受噪声和模糊的干扰很大。

小波变换的模极大值点对应于信号的突变点，在二维情况下，小波变换适用于检测图像的局部奇异性，故可通过检测模极大值点来确定图像的边缘。图像边缘和噪声在不同尺度上具有不同的特性，因此在不同的尺度上检测到的边缘在定位精度与抗噪性能上是互补的。在大尺度上，边缘比较稳定，对噪声不敏感，但由于采样移位的影响，使得边缘的定位精度较差；在小尺度上，边缘细节信息比较丰富，边缘定位精度较高，但对噪声比较敏感。因此，在多尺度边缘提取中，应发挥大、小尺度的优势，对各尺度上的边缘图像进行综合，以得到精确的单边像素宽的边缘。详细论述见 9.1 节。

4. 小波分析在数字水印中的应用

数字水印(digital watermarking)是一种新的有效的数字产品版权保护和数据安全维护技术，它是一种十分贴近实际应用的信息隐藏技术。以图像为载体对象的水印技术是当前研究的热点，由于小波分析在图像处理中所体现出的优势，目前已经有学者将小波分析用在了图像水印技术中。使用小波域水印方法的优点与在 JPEG 中使用小波是类似的，主要表现在以下三个方面[94]：一是可以保证在“JPEG-2000”有损压缩下水印不会被去除；二是可以将图像编码研究中关于视觉特性的研究成果用于水印技术；三是有可能提供在压缩域中直接嵌入水印的方法。除此之外，小波的多分辨率分析与人眼视觉特性是一致的，这对根据 HVS 选择适当的水印嵌入位置和嵌入强度有很大的帮助。

和其他变换域的水印技术一样，小波变换域水印也分为水印嵌入、提取(检测)两部分，小波水印的嵌入和提取都是在小波域中进行的。在此过程中，小波的类型、水印的选取、水印嵌入的强度以及水印嵌入的位置都会影响到水印系统的性能，包括水印的鲁棒性

和视觉可见性。

5. 小波分析在遥感影像融合中的应用[94]

基于小波分析的影像融合技术是目前遥感影像融合研究的主流。在低分辨率的多光谱影像和高分辨率的全色影像的融合中，要求充分利用多光谱影像的光谱信息与全色影像的细节信息，使融合后的多光谱图像具有较高的空间细节表现能力，同时较好地保持原始多光谱图像的光谱特性。传统的小波分析融合方法是在小波变换域中，用高空间分辨率的全色图像的细节分量替代低空间分辨率的多光谱图像的细节分量，然后对多光谱图像的小波系数进行小波逆变换，得到融合的多光谱图像。在小波变换域中，细节系数幅值较大的位置对应于灰度变化，即对应于边缘、突变点等显著特征，在融合中不做简单的替代，而是综合考虑两幅影像的显著特征，按一定的融合准则选取每一位置上的小波系数，再经过小波逆变换得到重构的融合影像。这就是基于特征的小波分析融合方法。

上面列出了小波分析在图像处理领域的一些应用，关于小波在模式识别、语音识别、量子物理、地震勘测、流体力学、电磁场、CT 成像、机器视觉、机器故障诊断与监控、分形、数值计算、微分方程计算等领域的应用，本书不再涉及，有兴趣的读者可参考相关文献[87～89,91,95]。

1.3 小波滤波问题、原理及方法

随着计算机应用技术的迅速发展，在数据分析和信号处理方面涌现出大量有趣的课题。从银河系天文学到分子光谱学，从 SAR 图像到医学图像，直至计算机视觉等领域，我们都需要从不完备的、直接或间接测量的，而且往往是被噪声污染的数据中恢复信号、曲线、图像、频谱或密度。这就使得利用各种信号处理理论、数学分析方法以及计算机技术对各种信息在一定准则下进行滤波、估计和分析的技术得以迅速发展[25～27]。

近年来，小波滤波这一概念不断见之于有关信号及图像处理研究的文献中[27～40]，这标志着一种新的信号滤波思想的出现。在早期的多尺度信号处理工作中，人们就已注意到信号和噪声在不同尺度上有不同的特征表现，并试图有效地利用这些特征，小波变换的出现为这一思想提供了一个自然而完美的工具，使信号图像的多尺度处理技术得到迅速发展。

20 世纪 90 年代初期，在一些公开出版的文献中开始出现有关小波去噪（denoising）的概念或术语，表示在小波域对噪声小波系数进行抑制，以达到去除噪声的目的[32,41]。Lu Jian 则多用术语降噪（noise reduction）[30,33,42,43]。直到 1994 年，Donoho 正式解释了“去噪”一词的含义，指的是通过对小波系数阈值化处理基本上可去除所有的噪声，在均方差意义下取得最优的去噪效果[37]。事实上，随着小波去噪方法应用的推广和深入，考虑到具体应用目的的不同，目前用得较多的术语还有估计（estimation）、回归（regression）、平滑（smoothing）和滤波（filtering）等，在国内外的文献中均有使用，且尚无统一标准。由于滤波的含义更为广泛，并且在小波分析中，除去噪外，还有如平滑、锐化和保留信号特征等功能，同时考虑到小波分析所具有的时频分析功能，作者认为“小波滤波”更能全面反映

问题。

基于小波分析和子带分解的边缘检测与滤除噪声的方法最早是由 Lu Jian 和 Mallat 几乎同时提出的[32,44]。Witkin 首先引入了利用尺度空间相关性来对信号滤波的思想[28]。对含噪信号经过子带分解后，从粗尺度到细尺度逐步确定信号的主要边缘，最终从噪声背景中得到真实信号。1992 年，Mallat 等人提出了基于信号奇异性（singularity）的信号和图像多尺度边缘表示法，利用 Lipschitz 指数在多尺度上对信号和图像及噪声的数学特性进行描述，并模极大值重构滤波方法[32,45]。Lu Jian 等人直接将小波变换理论与传统的多尺度信号处理方法相结合，给出了一种性能优良的小波滤波方法，并在医学图像处理领域得到应用[30,33,42]。Rosenfield 曾指出，在进行数字图像处理时，直接将相邻频带上的数据相乘，可以准确地定位信号边缘[46]。基于上述思想，Xu 提出了基于信号尺度间相关性的空域相关滤波算法（spatially selective noise filtration，SSNF）[35]。随后，斯坦福大学以 Donoho 为首的一个学术群体另辟蹊径，提出了小波域阈值滤波算法，取得了大量的理论及应用成果[25,37,47～50]。与此同时，Krim 等人运用最小描述长度（minimum description length，MDL）准则，得到了相同的阈值[51]。P. Moulin在研究谱估计问题时，得到了与阈值滤波算法相类似的结果[52]。不同之处是P. Moulin先设置了一个虚警门限，然后据此来选择阈值，使算法对信号成分和噪声成分错判的概率低于给定的门限。Moulin 将该方法用于谱估计，研究表明，这一方法具有相当的灵活性。此外，还有基于极大验后概率 MAP 的自适应收缩法等，都丰富了小波滤波的内容。

小波域滤波可按以下几个特征进行分类[53]：

（1）按所采用的变换方法分为进行基变换（快速小波变换、ridgelets、局部余弦基等）、利用框架来变换（非抽取小波变换）和通过选取最优基来变换（如小波包、自适应提升变换）。

（2）按所利用的小波性质分为能量集中性（阈值滤波算法）、多分辨率特性（尺度间相关性、树结构算法）、尺度间相关性（使用 Markov 随机场模型）和时间频率定位特性。

（3）按使用的模型分为 Bayesian 方法（全 Bayes 模型或经验 Bayes 模型）、非 Bayesian 方法（逼近理论、函数集合、平滑方法）等。

总而言之，近年来有关小波滤波的文献很多，并取得了不少的成果。作者一直从事小波滤波的研究，进行了大量深入的研究工作，也取得了一定的成果[96～115]。下面首先对小波滤波方法的一些相关问题作一概述，具体细节参考书中各章节。

1.3.1 基本原理

小波变换具有以下特性：

（1）时频局部化特性。小波变换可在时间轴上准确定位信号的突变点。

（2）多分辨率特性。小波变换可在不同尺度上描述信号的局部特征，如边缘、尖峰、断点等。

（3）解相关特性。小波变换可以对信号解相关，使信号的能量集中于少数几个小波系数上，而噪声能量分布于大部分小波系数上。

（4）选基灵活性。由于小波变换可以灵活选择小波基，从而可针对不同的应用对象

选用不同的小波函数。

性质(1)与(2)决定了小波滤波方法与传统方法相比,具有独特的优势——能够在去除噪声的同时,很好地保留信号的突变部分或图像的边缘[54]。

信号和噪声在小波域中有不同的性态表现,它们的小波系数幅值随尺度变化的趋势不同。随着尺度的增加,噪声系数的幅值很快衰减为零,而真实信号系数的幅值基本不变。在此基础上本书给出如下描述:小波滤波,就是利用具体问题的先验知识,根据信号和噪声的小波系数在不同尺度上具有不同性质的机理,构造相应规则,在小波域采用其他数学方法对含噪信号的小波系数进行处理。处理的实质在于减小甚至完全剔除由噪声产生的系数,同时最大限度地保留真实信号的系数,最后由经过处理的小波系数重构原信号,得到真实信号的最优估计。"最优"的精确定义依赖于应用要求。

小波滤波的特点可概括为:首先,它不是平滑(尽管许多文献用同义词 smoothing 来代替 denoising)。平滑是去除高频信息而保留低频信息;而小波滤波是要试图去除所有噪声,保留所有信号,并不考虑它们的频率范围。其次,它是在小波变换域对小波系数进行非线性处理。第三,滤波过程一般由三个步骤来完成:①小波变换;②对小波系数进行非线性处理,以滤除噪声;③小波逆变换。其中步骤②是本书讨论的重点,该非线性处理过程可以是 Mallat 提出的模极大值处理算法,或是由 Xu 等提出的空域相关滤波算法,也可以是由 Donoho 提出的阈值滤波算法。由于要对小波系数进行非线性处理,所以该过程不同于其他线性的滤波方法。最后,小波滤波是一种非参数方法。在参数方法中,必须要对一个特定的先验模型进行参数估计。例如,最常见的用最小二乘法来估计模型 $y=ax+b$ 中的参数 a 和 b。显然,小波滤波方法不同于上述的参数方法[55]。

1.3.2 小波滤波的数学模型及基本方法

对于乘性噪声的去除,虽有一些变换为加性噪声处理的方法,但总体上讲比较复杂,往往要针对具体问题方可得到较好的处理结果。限于篇幅,本书不予展开讨论。加性噪声研究具有普遍性,本书主要讨论具有加性噪声的数据处理和信号恢复方法。另外,小波滤波方法对噪声性质还有三个基本假设:①噪声经小波变换后大多数小波系数为零或近似为零;②噪声均匀地分布在所有系数中;③噪声水平不是太高[25]。

假设观测数据

$$f_i = g_i + \varepsilon_i, i = 1,2,\cdots,N(N = 2^n) \tag{1-3-1}$$

由真实信号 g_i 和加性噪声 ε_i 组成,其向量形式表示为 $\boldsymbol{f}=\boldsymbol{g}+\boldsymbol{\varepsilon}$。在理想情况下,$\varepsilon_i$ 为①正态噪声;②不相关的噪声;③方差为常量。当然实际情况并非如此,因此每种假设条件都可以适当放宽,以满足实际应用的需要。我们的目标是在观察到 f 的前提下,对 g 进行估计。

设对观测数据经过小波变换后,得

$$\boldsymbol{w} = \boldsymbol{\theta} + \boldsymbol{\eta} \tag{1-3-2}$$

设 $W(\cdot)$和 $W^{-1}(\cdot)$分别表示小波变换和逆变换算子,前述滤波过程的三个步骤可描述为

$$\boldsymbol{w} = W(\boldsymbol{f}), \boldsymbol{w}_t = D(\boldsymbol{w},t), \hat{\boldsymbol{g}} = W^{-1}(\boldsymbol{w}_t) \tag{1-3-3}$$

其中，$D(\cdot,\cdot)$为非线性滤波算子，它是滤波问题的核心。显然，这种原理性的归纳并没有涉及到算子$W(\cdot)$或$D(\cdot,\cdot)$是如何作用到信号上的，也没有涉及到它们的选取方法以及对滤波效果的影响。通过选择不同的$W(\cdot)$和$D(\cdot,\cdot)$，可以得到多种不同的小波滤波方法。

小波滤波的核心是在上述第二个步骤中按照一定的准则对小波系数进行修改，以在不损失过多信号的前提下，达到降低或去除噪声的目的。目前存在的方法主要可分为贝叶斯法和非贝叶斯法，其中贝叶斯法在4.1节讨论。非贝叶斯法大致可分为三种：①空域相关滤波，主要利用信号小波系数在各尺度间具有相关性来滤波；②模极大值重构滤波，主要根据信号和噪声小波系数随尺度增大具有不同的变化规律来滤波；③小波域阈值滤波，主要依据幅值较大的系数由重要信号产生这一基本假设来滤波。在第4章将对上述方法的原理和优缺点进行详细的分析比较。

1.3.3　小波域阈值滤波的基本问题

在小波域阈值滤波算法中，阈值函数选取和阈值确定是两个最基本的问题。这是本节要讨论的内容。

1. 阈值函数的选取

阈值函数体现了对小波系数的不同处理策略，主要分为软阈值函数、硬阈值函数和半软阈值函数。它们的基本思想都是去除小幅值的系数，对幅值较大的系数进行收缩或保留[25,37,47～50,56]。

Bruce 和 Gao 于1995年在本章参考文献57中得出结论：硬阈值法往往使得滤波结果具有较大的方差（主要因为滤波后的不连续性），而软阈值法使滤波结果有较大的偏差（主要因为其对所有大于阈值的系数共同做了收缩）。总的来说，硬阈值法可以很好地保留信号或图像的边缘等局部特征，但滤波结果会出现伪 Gibbs 现象等视觉失真，而软阈值法的滤波结果则相对平滑得多，但可能会造成边缘模糊等失真现象。为了克服软阈值法和硬阈值法的缺点，Gao Hong-Ye 提出了另一种阈值函数，即半软阈值（semisoft shrinkage）函数，并在此基础上推导出了基于半软阈值法的 Minimax 阈值[56,58]。这种方法是软阈值法和硬阈值法的一种折衷形式。它不仅保留了较大的系数，而且具有连续性。然而在半软阈值法中，需要确定两个阈值，增加了算法的复杂度（详见4.4节）。Gao Hong-Ye 随后又提出了用 Garrote 函数作为阈值函数，并证明了上述的各种阈值函数得到的滤波结果是渐近相等的[59]。另外，Zhang 从选取最优阈值的角度提出了一种具有高阶连续导数的阈值函数[60,61]，可对 SURE（Stein's unbiased risk estimator）函数利用梯度下降迭代法进行优化搜索。由于采用高阶可导的阈值函数，其滤波效果得到了改善。除了上述的阈值函数，还有很多对基本的硬阈值和软阈值函数的改进，可参考本章参考文献83、84。

以上所讨论的阈值函数，可统称为显式阈值函数，它们都是在假定真实信号为确定性信号的前提下进行的。另一类隐式的阈值函数是基于贝叶斯模型得到的[62～68]。贝叶斯方法认为真实信号为随机信号。首先假定真实信号的小波系数服从某一先验分布，然后通过贝叶斯准则计算信号系数的后验密度，最后计算后验均值作为信号小波系数的后验

估计，在此基础上得到一种隐式的阈值化方法。通过选取不同的先验分布密度，可以得到不同的阈值收缩函数。这时，先验分布的选择是一个关键问题，而且超参数(hyperparameter)的计算非常复杂。另一种相关的方法是由 Crouse 在 1997 年提出的利用隐马尔可夫模型来估计系数[69,70]。这类隐式阈值函数的选取需要坚实的统计学基础，本书不作深入讨论。

2. 阈值的确定

在小波域阈值滤波中，阈值的选取直接影响滤波效果。目前有大量的文献提出了各种各样确定阈值的方法[47,48,71~79]，其中主要有通用阈值法(universal method)、极小化风险阈值法、假设检验法和 BayesShrink 阈值法等。下面分别对它们进行简要地概括与比较，具体讨论详见第 6 章。

(1) 通用阈值法

这是 Donoho 和 Johnstone 在高斯噪声模型下，应用多维独立正态变量决策理论，得出的阈值[37]。他们发现当维数趋于无穷时，噪声系数的幅值大于阈值 $t=\sigma\sqrt{2\ln N}$ 的概率趋于零，其中 σ 为噪声标准差，N 为信号长度。通用阈值具有渐进的 Minimax 特性，然而由于侧重考虑滤波结果的平滑性，结果表现出较大的偏差。实际应用时需估计噪声方差，潘、张等从理论上得到了 σ^2 的一致无偏估计[96,100,105]。然而在软阈值法中应用该阈值进行滤波时，往往会过平滑掉真实信号。潘、张等人提出了自适应阈值 $t=c\sigma$，其中系数 c 可通过极小化均方差函数来自适应地选取以得到最优阈值。另外，该方法还验证了在一般情况下，基于正态分布“3σ 准则”的阈值 $t=c\sigma$，$c\in[3.0,4.0]$ 是合理的[101,106]。

(2) 极小化风险阈值法

均方差(mean square error, MSE)函数的期望称为风险，极小化风险阈值即 MSE 意义下的最优阈值。均方差函数描述了滤波结果与真实信号在均方意义下的偏离程度，但是由于在实际应用中，真实信号是未知的，所以人们提出了许多方法来对 MSE 函数进行估计，这些方法主要有 SURE 法、交叉验证算法和广义交叉验证算法等。

① SURE 法

Donoho 和 Johnstone 提出通过极小化 SURE 准则函数来确定阈值，大数定理保证了 SURE 函数是均方差损失函数的无偏估计。所以，SURE 阈值也是最优阈值的无偏估计。

② 交叉验证(cross-validation，CV)算法

G. P. Nason 提出的 CV 算法是一种基于均方差准则确定最优阈值的统计方法[71]。CV 算法可以根据实际问题的需要自适应地选取不同的估计准则，如估计误差可以用 L_∞ 或 L_1 范数等来衡量，以取得更好的滤波效果。当 CV 算法用于相关噪声滤波时，即使采用基于尺度自适应的阈值，其滤波效果仍不理想。

③ 广义交叉验证 (generalized cross-validation, GCV)算法[72,73]

另外一些学者提出的 GCV 算法，是以 SURE 为基础的，其性能远优于 Nason 的 CV 算法。GCV 是有偏的，其偏差为 σ^2，但其得到的最优阈值是无偏的，且不需要估计噪声方差，故 GCV 算法应用很广泛。

(3) 多假设检验法

阈值处理过程可看作是一个多假设检验过程[79]。Abramovich 和 Benjamini 引入了错判率(false discovery rate,FDR)的概念,它指的是经过阈值化处理后,错判的非零系数个数与所保留的非零系数总数比值的期望。在满足给定的 FDR 上界的前提下,使所保留的系数个数达到极大值的阈值,即为所要求取的最优阈值。FDR 滤波方法是在给出错判率的前提下,导出阈值。此方法与通用阈值法相比有一定灵活性,而且它可以从本质上解释通用阈值依赖于信号长度这个现象。当然 FDR 滤波方法也有其局限性,其一是如何给定错判率,其二是错判率只考虑了将噪声误判为信号的情况,却没有考虑将信号误判为噪声的错判率,也就是说着重考虑了滤波结果的平滑性问题。

(4) BayesShrink 阈值法

这种阈值主要是针对二维图像提出的。在假设自然图像的小波系数先验分布为广义高斯分布的前提下[74],通过极小化贝叶斯风险,得到阈值 $T_{\mathrm{B}}=\sigma^2/\sigma_X$。其中,$\sigma^2$ 为噪声方差,σ_X^2 为真实信号▷◁的方差。

以上阈值中,通用阈值计算简单,故得到广泛应用。但是在 N 较大时显得过大,往往过平滑掉信号细节,从而导致较大的重构误差;FDR 阈值虽与信号长度无关,但由于仅考虑了错误拒绝却没有考虑错误接受假设"小波系数显著为零"的情况,因而滤波结果偏差较大,这在图像处理中表现为滤波后图像变模糊;极小化风险阈值是使重构误差极小的阈值,因而 SURE 阈值和 GCV 阈值往往能获得较为满意的滤波效果,但结果中有时会含有"毛刺";实验表明,BayesShrink 阈值的滤波效果接近于极小化 MSE 阈值。

若阈值在整个滤波过程中固定不变,即对所有的小波系数均采用相等的阈值进行收缩,可将其称为全局阈值,而自适应阈值是根据小波系数的局部特征来对其进行阈值化处理,最简单的自适应阈值是可以适应于不同尺度上信号和噪声具体特征的阈值,也就是基于尺度自适应(scale-adaptivity)的阈值。另外,还有基于小波系数尺度内相关性和尺度间相关性的自适应阈值等,这就涉及到了对小波系数进行建模的问题。

1.3.4 小波域系数模型

在小波域对小波系数建模很重要。只有建立了准确的小波系数模型,才会有更有效的小波滤波方法。这一点对图像小波系数建模显得尤其重要。小波系数模型主要可分为基于尺度内相关性的尺度内模型、基于尺度间相关性的尺度间模型和混合模型[80,81]。其中,尺度内模型主要考虑同一尺度内小波系数的统计特性以及相邻系数间的关系;尺度间模型是考虑不同尺度上小波系数相关性的模型;而混合模型则综合考虑了尺度内和尺度间小波系数的关系。详见 7.5 节。

1.4 本书的内容安排

全书共分 9 章,除第 1 章引言对全书内容作概述外,其余各章主要内容如下:

第 2 章从傅里叶变换开始,沿着傅里叶变换→短时傅里叶变换→小波变换的发展轨迹,从物理直观的角度对其逐一进行介绍,引出了小波变换的概念。然后对小波变换的基本理论进行了比较详细的叙述。最后介绍了小波理论发展的两个方向——向量小波和提