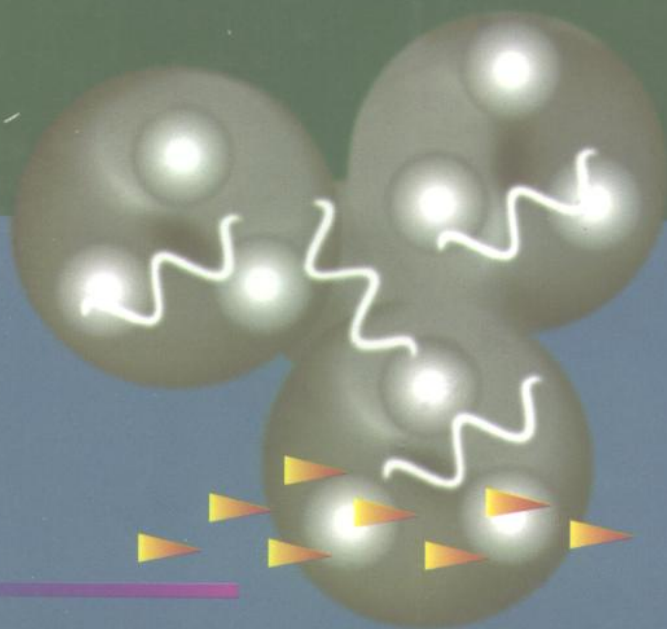




神经网络与信号处理系统

——有限精度设计理论

何明一 保 铮 著



神经网络与信号处理系统 ——有限精度设计理论

何明一 保 铮 著

国家自然科学基金资助项目

国家基础研究重大关键项目计划资助课题

西北工业大学出版基金资助出版

西北工业大学出版社

1998年11月 西安

(陕)新登字 009 号

图书在版编目(CIP)数据

神经网络与信号处理系统:有限精度设计理论/何明一,保铮
著.—西安:西北工业大学出版社,1998.11

ISBN 7-5612-1049-3

I. 神… II. ①何… ②保… III. ①神经网络—精度—设计
②信号处理—精度—设计 IV. TN911.7

中国版本图书馆 CIP 数据核字(98)第 13649 号

神经网络与信号处理系统——有限精度设计理论

何明一 保 铮 著

责任编辑 雷 鹏 张近乐

责任校对 齐随印

*

©1998 西北工业大学出版社出版发行

(邮编:710072 西安市友谊西路 127 号 电话:8493844)

全国各地新华书店经销

西北工业大学出版社印刷厂印装

ISBN 7-5612-1049-3/TP·151

*

开本:850 毫米×1 168 毫米 1/32

印张:12.875 字数:310 千字

1998 年 11 月第 1 版

1998 年 11 月第 1 次印刷

印数:1—2 000 册

定价:29.00 元

购买本社出版的图书,如有缺页、错页的,本社发行部负责调换。

前 言

近年来,信号处理与电路系统的理论与方法得到了迅速发展。无论是电子系统的模拟信号处理、数字信号处理,还是以神经网络为重要标志的智能信号处理,在模型、理论、应用系统等方面都取得了实质性的突破,并在处理各种自然的与人工的实际问题中得到了许多应用。

作为应用科学,在信号处理与电路系统研究中,模型是基础,应用是目的,理论分析则是由模型通向应用的桥梁,而设计理论与实现技术是使信号处理系统和神经网络真正得到实际应用的关键。从数学和系统模拟的观点来看,很多系统都是完全等价的,也就是说一个理想系统的实现方案可以有多种。但是,在用物理元器件、模块或子系统实现系统时,系统的性能与系统的实现结构有着密切的关系,理想条件下的相似系统变得并不相似,甚至大相径庭。其原因在于:各种组件的有限精度效应,例如,模拟元器件的有限精度(特别是温漂和老化等的影响)和数字系统的有限字长效应;系统精度的提高又受到成本的制约,例如,VLSI 系统中,系统成本的增加与数字系统字长的平方甚至立方成正比;复杂性也是必须要考虑的问题,在 VLSI 中,一块芯片能容纳的电路元件(例如门电路)是有限的。另外,高精度的表示与处理需要复杂得多的内部电路、接口和外围电路,在要求速度高、体积小、功耗低的场合,例如,弹载、机载、星载和人体内脏携带的信号处理系统等场合,系统的体积和功耗受到极大的限制。总之,建立在无限精度设计、有限精度工程实现的传统模式应该发展成为有限精度设计、有限精度工程实现的模式。因此,有限精度设计与实现理论问题的重

要性就不言而喻了。

有限精度设计理论这一课题早在模拟信号处理系统设计时就提出来了,数字系统设计中转化成有限字长效应分析。信号处理系统的有限精度设计与实现理论的研究难度是相当大的。70年代,国际著名信号处理学者、美国麻省理工学院的 A. V. Oppenheim 教授曾预言,优化灵敏度的系统设计理论要取得大的进展在短期内几乎是不可能的,因而多年来都是采用先按其它设计要求设计出多种实现结构,然后分别分析或检验它们的有限精度效应,并进行择优。在神经网络系统设计中,主要采用模拟电路和模数混合电路,有限精度理论变得更为重要。可以说,神经网络发展到今天,低成本、高性能的设计与实现理论成为神经网络进一步发展和广泛应用的关键问题之一。近几年,在世界上的许多有关研究机构 and 高等院校中都正在从事这一领域的研究。

我们在神经网络与信号处理系统的有限精度设计与实现理论及其应用这一领域的研究开始于80年代中期。这些研究先后得到了国家和有关部委科学基金项目的支持,其阶段成果有的已经在国际、国内学术刊物与会议上发表并受到国内、外同行专家的好评。有的结果以论文形式发表后,先后被国际权威索引 SCI, ISTP, EI, INSPECT, SA, Key Abstract 等收录。此外,有关的阶段成果已经获得国家教委、中国航空工业总公司和陕西省等颁发的科技进步奖,有的理论成果已经应用于 X 射线多光谱行李成像和某飞行器信号处理系统中,取得了较好的应用效果。

虽然信号处理系统和神经网络的有限精度设计理论和方法还不够成熟,但鉴于这一课题在理论上和应用中都有重要的意义以及还未见到过论述有限精度设计理论的专著公开出版,我们认为及时地把这一课题已经取得的成果加以收集和系统的整理是有意义的,因为这有利于指导我们今后的研究工作。另外,我们希望通过对这一系统的整理能引起更多的人重视神经网络与信号处理系

统的有限精度设计与实现理论,从而使这一理论在各个方面得到进一步发展和更广泛的应用。当然,将这一理论广泛应用于实际还需要从理论上和技术上解决一系列重要问题,例如,其中一个重要的问题就是神经网络有限精度实现对网络推广性能的影响,这也是我们正在研究的课题之一。本书正是在这样一种形势下完成的。本书论述神经网络与信号处理系统的有限精度设计与实现理论及其应用,主要是基于近 10 年来这一领域国内、外的研究进展及成果,尤其是包括了我们在完成科研项目时取得的一些成果和研究中的体会与心得。本书编写计划是由我们共同讨论的,由何明一完成初稿后再共同讨论定稿。

无论是系统还是网络都有线性与非线性之分,也有连续时间和离散时间之分。我们在本书中以神经网络为非线性系统的代表,以输入-输出表示或状态空间表示的信号处理系统为线性系统代表(虽然实际上也可以是非线性的,但我们在本书中仅限于线性系统)。对于线性的信号处理系统,我们又按连续时间和离散时间系统分别加以讨论。这样的考虑和安排,是为了建立这三类系统(或网络)有限精度设计的统一理论,使研究有可能系统、全面,使结果可能更具有普遍意义。

我们在内容组织与安排上有如下考虑:

(1) 模拟信号处理系统方面重点论述低灵敏度模拟滤波器的设计理论、结构、准则、方法和步骤。

(2) 数字信号处理系统部分分别按输入-输出描述和状态空间描述,重点研究有限字长效应的设计理论和方法。

(3) 神经网络部分重点论述神经网络的有限精度效应分析与有限精度设计理论。神经网络有限精度设计理论是本书重点,占大半篇幅,内容包括:① 建立了无限与有限精度神经网络的统计模型;② 讨论了分段线性神经网络的量化效应;③ 分别按一元和多元函数线性近似的方法分析了多层前向神经网络的权灵敏度;④

用统计模型研究了网络的范数灵敏度和噪信比,并给出了降低灵敏度与噪信比的设计方法;⑤进一步放宽条件研究了前向神经网络的误差界,并进行了数值仿真;⑥从设计的观点研究了有限精度神经网络的学习算法(包括归一化学习、减弱误饱和效应法、导数修正法、纠错法、基于灵敏度和凸值的修枝算法、权值衰减法、权扰动法等);⑦研究了映射能力与权值范围的关系及其数值仿真,揭示了大权值与小权值在网络中的不同作用;⑧最后,研究了有限精度神经网络的数值仿真和基于 DSP 的神经网络有限精度实现系统及其应用。

我们从事的与本书有关的研究工作先后得到了国家自然科学基金项目(批准号 69171025, 69572032)、国家基础研究重大关键项目计划(即攀登计划,项目资助号 NSF 9207)、航空科学基金、国家教委优秀年轻教师基金和国际合作项目的资助,本书也是这些项目的成果之一。

本书第一作者在科研工作中,得到国家攀登计划(神经网络)首席科学家何振亚教授、中国工程院院士吴佑寿教授、西北工业大学校长戴冠中教授的热情关心、支持和帮助,同前任国际神经网络学会会长 P. J. Werbos 博士和 H. Szu 教授等进行过有益的讨论,在此,谨致以最诚挚的感谢。

在研究过程中,冯燕副教授、杨湘玉同学等参与了部分实验工作。在写作过程中,杨湘玉、邴军、夏建涛、陈立和王敏同学帮助抄写了部分书稿。此书的出版得到了西北工业大学出版社的有关领导和同志的关心和支持。对此,作者表示衷心的感谢。

由于本书论述的主题正在迅速发展之中,限于作者的水平,在取材、组织安排和论证分析等方面难免会有不足之处,殷切希望各位专家和广大读者批评指正。

著 者

1998 年 3 月于西安

目 录

第 1 章 绪论	1
1.1 无限精度与有限精度	1
1.2 有限精度设计理论研究的意义	2
1.3 有限精度设计理论研究的背景与动机	3
1.3.1 模拟系统	3
1.3.2 数字系统	4
1.3.3 神经网络系统	5
1.4 有限精度设计理论研究的状况与趋势	6
1.5 有限精度设计理论研究的主要内容	7
1.6 本书内容安排	9
参考文献	11
第 2 章 信号处理系统有限精度设计理论基础	14
2.1 信号及其分类	14
2.1.1 信号与映射	14
2.1.2 连续信号与离散信号	14
2.1.3 确定性信号与随机信号	15
2.1.4 确定性能量有限信号	15
2.2 系统及其分类	16
2.2.1 连续系统与模拟系统	16
2.2.2 离散系统与数字系统	18
2.2.3 神经计算系统	19

2.3 有关灵敏度的概念和定理	23
2.3.1 Telegen 定理	23
2.3.2 互易的和交相互易的网络	25
2.3.3 灵敏度	26
2.4 本章小结	30
参考文献	31
第3章 神经网络基础	32
3.1 概述	32
3.2 神经元模型	32
3.2.1 神经元结构	32
3.2.2 神经元的运算	33
3.2.3 神经元的活动函数	35
3.2.4 神经元的内部反馈	36
3.3 神经网络的基本结构	38
3.3.1 网络功能层次结构	38
3.3.2 网络拓扑结构	39
3.4 神经网络的学习与功能神经网络	43
3.4.1 学习方法分类	43
3.4.2 误差后向传播网络	45
3.4.3 反馈联想网络	48
3.4.4 竞争网络	52
3.4.5 自适应共振网络	56
3.4.6 其它网络	57
3.5 神经网络计算的统一处理	58
3.6 一般多层前向网络与双并联前向网络	58
3.6.1 网络结构	58
3.6.2 工作原理	61

3.6.3 映射功能与误差曲面	63
3.6.4 学习	64
3.6.5 双并联前向网络的误差后向传播学习算法	73
3.7 本章小结	85
参考文献	86
第4章 低灵敏度模拟滤波器设计理论	88
4.1 概述	88
4.2 模拟滤波器的实现结构	88
4.2.1 直接实现	89
4.2.2 串联实现	89
4.2.3 并联实现	91
4.3 零点、极点与参数灵敏度	92
4.3.1 系统的零点与极点	93
4.3.2 系统极点对系数的灵敏度	93
4.3.3 系统零点对系数的灵敏度	95
4.3.4 ω_j, Q_j 对参数的灵敏度	96
4.4 可调滤波网络的双二次模型	98
4.4.1 分析模型	99
4.4.2 构造模型	100
4.5 可调双二阶网络的灵敏度分析	102
4.5.1 一组新的四端口参数	102
4.5.2 传递函数	104
4.5.3 可调双二阶网络的灵敏度分析与 最低灵敏度准则	106
4.6 并联与串联网络的灵敏度	115
4.6.1 并联网络的灵敏度	115
4.6.2 串联网络的灵敏度	117

4.7 可调最低灵敏度网络的设计与实现	118
4.7.1 低灵敏度二阶网络的设计	118
4.7.2 基于运算放大器的设计实例	122
4.7.3 基于可调跨导器件的设计	125
4.8 SPICE 集成电路软件仿真	128
4.9 本章小结	131
参考文献	132
第5章 数字信号处理系统有限精度设计理论	133
5.1 概述	133
5.2 信号的二进制表示及其误差	136
5.2.1 定点表示	136
5.2.2 浮点表示	138
5.2.3 舍入与截尾误差	139
5.3 模数转换器的量化效应	143
5.3.1 模数转换器的取样与量化	143
5.3.2 量化的统计表示	144
5.3.3 等效的输入量化误差	147
5.4 I/O 表示的数字滤波器对系数的有限精度效应	148
5.4.1 极点、零点对系数变化的灵敏度	149
5.4.2 零、极点网格图	152
5.5 定点运算数字滤波器的有限精度效应	157
5.5.1 概述	157
5.5.2 IIR 滤波器中的零输入极限环	158
5.5.3 舍入误差的统计分析	162
5.5.4 溢出误差的统计分析	180
5.6 浮点运算数字滤波器的有限精度效应	186
5.7 DFT 计算中的有限精度效应	190

5.7.1 直接 DFT 计算过程中量化效应分析	191
5.7.2 定点 FFT 中的有限精度效应	194
5.7.3 浮点 FFT 中的有限精度效应	202
5.7.4 系数量化对 FFT 的影响	205
5.8 线性调频 Z 变换的有限精度效应	207
5.9 状态空间数字滤波器的低灵敏度设计	211
5.9.1 离散线性移不变系统的状态方程	211
5.9.2 状态方程系统的灵敏度	212
5.9.3 L_p 范数灵敏度与均方误差	214
5.9.4 $H(z)$ 对系数矩阵的 L_2 灵敏度	217
5.9.5 L_2 灵敏度的 Gramian 矩阵	222
5.9.6 L_2/L_1 灵敏度的改进	227
5.9.7 L_2 灵敏度模型的化简	230
5.10 本章小结	233
参考文献	236
第 6 章 无限与有限精度神经网络	238
6.1 概述	238
6.1.1 神经网络有限精度设计理论的研究动机	238
6.1.2 神经网络有限精度设计理论研究的 发展状况与趋势	239
6.2 神经网络有限精度统计模型	241
6.3 无限精度神经网络模型	243
6.4 有限精度神经网络模型	244
6.5 权矢量的统计模型	246
6.6 输入矢量的统计模型	247
6.7 权误差与输入误差的统计模型	247
6.8 分析中的其它考虑	248

6.9 本章小结	249
参考文献	250
第7章 分段线性神经网络的量化效应	253
7.1 概述	253
7.2 输入、权值的量化表示	253
7.3 隐层有限精度效应分析	255
7.4 一般层有限精度效应分析	261
7.5 输入编码的影响	262
7.6 双并联前向神经网络有限精度效应分析	264
7.7 信噪比分析	265
7.8 位数对各层的影响	267
7.9 本章小结	269
参考文献	269
第8章 多层前向神经网络的灵敏度分析	271
8.1 概述	271
8.2 多层前向神经网络权灵敏度分析(1) ——一元函数线性近似	271
8.2.1 非线性活动函数的处理	271
8.2.2 符号与意义	272
8.2.3 输出对权扰动的统计灵敏度定义	273
8.2.4 权扰动造成的输出误差	274
8.2.5 对加性权扰动的灵敏度	276
8.2.6 对乘性权扰动的灵敏度	278
8.2.7 对输入扰动的灵敏度	280
8.2.8 对单层权扰动的灵敏度	281
8.2.9 讨论	282

8.3 多层前向神经网络权灵敏度分析(2)	
——多元函数线性近似法	283
8.3.1 多元非线性函数线性化	283
8.3.2 网络对加性权扰动的灵敏度	285
8.3.3 网络对乘性权扰动的灵敏度	288
8.4 范数灵敏度	290
8.4.1 输出/输入导数矩阵	290
8.4.2 输出/输入导数矩阵的范数灵敏度	291
8.4.3 输出/权导数矩阵	293
8.4.4 输出对权导数矩阵的范数灵敏度	293
8.5 降低灵敏度的方法	295
8.5.1 降低权灵敏度	295
8.5.2 降低输入灵敏度	296
8.6 本章小结	296
参考文献	297
第9章 连续活动函数神经网络的噪信比	298
9.1 概述	298
9.2 单神经元的输出噪信比	298
9.2.1 单神经元的输出误差	299
9.2.2 单神经元输出误差的方差与输出噪信比	300
9.3 多层前向神经网络的输出噪信比	303
9.3.1 概述	303
9.3.2 第一层神经元的输出误差方差	304
9.3.3 第一层神经元的输出噪信比	306
9.3.4 噪信比增益函数	306
9.3.5 其它层神经元的噪信比	309
9.4 双并联前向神经网络的噪信比分析	311

9.5 本章小结	316
参考文献	317
第 10 章 前向神经网络的误差界分析	318
10.1 概述	318
10.2 若干假设	318
10.3 各层神经元总输入扰动的期望与方差界	320
10.3.1 神经元总输入扰动的期望界	320
10.3.2 神经元总输入扰动的方差界	321
10.3.3 输出状态的期望界	322
10.4 各层神经元输出的方差界	323
10.5 本章小结	328
参考文献	329
第 11 章 有限精度神经网络学习算法	330
11.1 概述	330
11.2 归一化学习	330
11.3 后向误差信号与误饱和	335
11.4 导数修正法	341
11.5 误差修正法	341
11.6 纠错法	346
11.7 基于灵敏度和凸值的修枝算法	346
11.8 权值衰减法	348
11.9 权扰动法	349
11.10 其它方法	350
11.11 本章小结	351
参考文献	351

第 12 章 多层前向网络映射能力与权值范围的关系

及其仿真	354
12.1 多层前向网络的逼近性能	354
12.1.1 多层前向网络的逼近特性	354
12.1.2 多层前向网络逼近算法特点	355
12.1.3 三层前向网络逼近映射的描述性说明	356
12.2 权值范围与映射性能	357
12.2.1 大权值与小权值在网络中的不同作用	357
12.2.2 三层网络中两层权值的联系	358
12.3 权灵敏度分析	360
12.3.1 选择表示灵敏度的量	361
12.3.2 三层前向网络权灵敏度分析	362
12.3.3 数值积分	365
12.3.4 权值大小对隐层输出误差的影响	367
12.4 隐层-输出层误差分析	368
12.5 灵敏度分析与仿真	369
12.6 本章小结	372
参考文献	373

第 13 章 有限精度神经网络的实现方法与仿真实验 374

13.1 概述	374
13.2 有限精度函数	377
13.3 输入有限精度表示及其影响	378
13.4 权值有限精度的影响	382
13.5 节点输出值与非线性函数的有限精度效应	383
13.6 权值修正量的有限精度影响	384
13.7 全量化实现	386
13.8 限制权值范围与变量增益	386

13.9 有限精度神经网络的 DSP 实现与应用简介	388
13.10 本章小结	389
参考文献	390
第 14 章 总结与展望	392
14.1 回顾与一般性结论	392
14.2 未来研究的发展方向	394