

分类号_____

UDC_____

学 号_____

重慶大學

硕士学位论文

论文题目：基于小波的图像压缩编码算法研究

论文作者：王文涛

指导教师姓名、
职称、工作单位：朱冰莲 副教授 重庆大学

申请学位级别：硕士 专业名称：信号与信息处理

论文提交日期：2005 年 4 月 25 日 答辩日期：2005 年 5 月 31 日

学位授予单位：重庆大学 授位日期：2005 年 月 日

答辩委员会主席：杨济安 教授

论文评阅人：吴乐华 教授 石 锐 副教授

2005 年 5 月 31 日

重庆大学硕士学位论文

基于小波的图像压缩编码算法研究



硕士研究生：王文涛

指导教师：朱冰莲 副教授

学科专业：信号与信息处理

重庆大学通信工程学院

二〇〇五年五月

Master Thesis of Chongqing University

Research for image compression algorithm
based on wavelet transform



Master Candidate: WenTao Wang

Supervisor Associate-Prof. Binglian Zhu

Major: Signal and Information Processing

College of Communication Engineering

Chongqing University

May,2005

摘 要

图像作为传递信息的媒体和手段是十分重要的。但是数字图像的数据量非常大,需要很多的比特数,因而需要对图像进行压缩。所谓图像压缩就是在保证一定的图像质量和满足一定要求的前提下,减少图像原始数据量的一种过程。针对图像中存在的相关冗余信息,采用某种方法去除多余成分,就能实现图像的压缩。尽管数据存储技术不断发展、信道传输带宽不断加宽,但人们对于压缩图像数据以节省数据存储空间和提高信道利用率的需求仍在增长。

传统的图像压缩采用 DCT 变换编码会出现“方块效应”,同时压缩的效率也不高,小波图像编码是近年来随着小波分析理论的发展而提出的一种具有很好发展前景的图像编码方法。小波分析具有时—频分析、多分辨率分析等优点,易与人眼视觉特性相结合,可获得较好的压缩效果,特别适合于图像这一类信号的处理,所以小波变换的方法受到人们的高度重视,出现了各种基于小波变换的图像编码方法。

本文主要研究了基于整数小波变换的图像编码。具体研究了提升方法、零树编码算法。提升方法属于第二代构造小波的新方法,既继承了第一代小波的特性,又有不依赖于傅里叶变换,同时便于构造整数小波的特性。本文对提升方法的原理和其中的关键技术作了详细的阐述和讨论,并选择了 5/3 和 9/7-M 两种小波用提升方法实现从整数到整数的小波,为后面进一步研究系数编码做好准备。对小波系数编码的方法是实现图像压缩的关键技术,不仅关系到图像的最终压缩效果、也关系到图像的恢复质量和编解码的时间消耗。本文详细阐述了基于零树思想的小波系数量化方法,并针对其存在的算法复杂、运算量大的缺点做出了相应改进,提出了改进的零树编码算法-----不完全扫描法。

最后,本文提出了一种基于整数小波变换和改进零树量化方法相结合的图像编码方案,并用 C/C++ 语言实现了相应的编/解码器。通过对标准测试图像进行压缩编码测试实验,证实该方法可以实现图像从完全无损到有损压缩,与基于传统小波变换及零树量化的方法相比,运算简洁,速度快,重构图像质量高,取得了令人满意的效果。

关键词: 小波变换, 零树编码, 提升算法, 图像压缩

ABSTRACT

The image as delivering medium and means of information are very important .But the data quantity of the numerical image is very big, needing a lot of ratios to especially count, as a result need to carry on the compression to the image .So-called image compression is at guarantee the certain image quantity and satisfy the premise of certain request under, a process of reducing the quantity of original data of image .Aiming at the existent and related redundancy information in the image s, adopting a certain method to do away with the surplus composition, can carry out the compression of the image. Though the saving technique of data develops continuously, the letter way delivers the bandwidth to widen continuously, people for the compression image data with economical data saving the space and exaltations believe the need of an utilization to increase still.

The traditional image compresses of adopting the DCT transformation to code would appear “a block effect”, the efficiency compress at the same time is also not high, a wavelet image coding is in recent years along with a development and analyzes of the wavelet theories to put forward a kind of way with very good development foreground of image code the method .the wavelet analysis have many advantages of time-frequency analysis, many differences analysis etc., easily combining together with the person's eye sense of vision characteristic, can get a good compression result, it is specially and suitably suit for compressing process of this signal, so the method of a wavelet transformation is high among the people value, and appearing various image coding methods according to a wavelet transformation .

This text mainly studies to code according to the image of a transformation of integral .Studied to rising the method, zero trees to code the calculate way in a specific way .Rising the method belongs to the new method of the wavelet of the next generation structure, since inherits the first generation characteristic of the wavelet, and then has not to be depend on Fourier transformation, at the same time it's easy to structure integral characteristic of the wavelet .This text elaborates and discusses towards rising principle and key technique in which of method to make detailedly , and chooses 5/3 and 9/7-M two kinds of wavelet of rising methods of realization from integral to the wavelet of the integral, for the behind further research coefficient codes to work well the preparation .That a coefficient code is a key technique that carries out

the image compression to the method , not only relating to the end compression result of the image, but also relating to the instauration quantity of the image and the time depletion of the decoding .This text is detailed to elaborate the method according to an measure of coefficient of zero the tree thoughts, and aimed at the weakness of complication of calculating way and operating that it exists and put forward a good improvement of zero tree code the calculate way- Scan the method incompletely.

Ending originally text, it puts forward a kind image coding methods according to combining together of a transformation of integral wavelet and improve zero measure of trees, and with the C/ C++ the language carries out the coding/decoding machine .with the method of carrying on compress to code the test the experiment to the standard test images, it confirms that that method can carry out the image transformation from completely compression of lossless to the compression of losing ,comparing with the method according to a wavelet transformation and zero amount of trees measure , it operates more simply and the speed is quicker, the composition image has quantity higher, obtains the decent result.

Keywords: wavelet transform , Zerotree Wavelet Encoding, Lifting Algorithm, image compression, lossy compression

目 录

中文摘要.....	I
英文摘要.....	II
1 绪 论.....	1
1.1 引 言.....	1
1.2 研究的背景及意义.....	1
1.3 小波图像压缩编码研究现状.....	2
1.4 本文主要内容及章节安排.....	4
2 图像压缩编码理论概述.....	5
2.1 图像压缩编码的基本原理.....	5
2.2 图像压缩编码技术的发展.....	5
2.3 图像编码的基本方法.....	6
2.3.1 熵编码.....	7
2.3.2 变换编码.....	8
2.3.3 预测编码.....	8
2.3.4 运动补偿编码.....	9
2.4 图像编码方法的评价标准.....	9
2.5 图像编码的国际标准.....	11
3 小波分析理论.....	15
3.1 引言.....	15
3.2 小波发展的历史背景.....	15
3.3 小波变换的基本理论.....	17
3.3.1 傅立叶变换及其存在的不足.....	17
3.3.2 小波的概念.....	17
3.3.3 连续小波变换.....	18
3.3.4 傅立叶变换与小波变换的比较.....	19
3.4 离散小波变换.....	20
3.5 多分辨率分析与 Mallat 算法.....	20
3.6 二维离散小波变换.....	22
3.7 提升算法.....	22
3.7.1 提升算法的过程.....	23
3.7.2 提升算法的实现.....	24

3.7.3 提升算法与传统小波变换的比较	25
3.7.4 基于提升方法的小波变换	26
4 基于嵌入式思想的小波系数编码	28
4.1 引言	28
4.2 嵌入式编码	28
4.2.1 嵌入式零树编码	29
4.2.2 零树编码的概念	30
4.2.3 零树编码算法原理	32
4.3 分层树集合分割算法 (SPIHT)	34
4.3.1 集合分类与分级量化	35
4.3.2 空间方向树	35
4.3.3 编码方法	36
4.4 嵌入式块集合分裂编码器 (SPECK)	38
4.4.1 块状集合分割	38
4.4.2 编码方法	39
4.5 现有算法的不足	40
5 基于整数小波变换的图像压缩编码方案	42
5.1 基于提升算法的整数小波变换	42
5.2 对嵌入式零树编码算法的改进——不完全扫描法	44
5.3 算术编码	46
5.4 测试图像的选取	49
6 实验结果及分析	52
6.1 无损压缩	52
6.2 有损压缩	53
6.3 编/解码时间	55
7 全文总结与展望	58
7.1 主要结论	58
7.2 后续研究工作的展望	58
致 谢	60
参考文献	61
附 录	65

1 绪 论

1.1 引 言

随着信息技术的发展和社会的不断进步,人类对信息的需求越来越丰富,图像作为信息最重要的载体之一,已经深入到人们的日常生活。图像信息对人类有着极其重要的意义,这主要表现在:

- 1、人类接收的信息约有 85%来自视觉,周围景物在视网膜上的映像是人类最有效、最直接,也是最重要的信息获取形式;
- 2、图像信息具有确切、直观、具体、生动、高效、广泛等等一系列的优点;
- 3、图像信息容量大,通过视觉获得的图像信息往往比通过其他感觉形式获取的信息具有更大的信息量,即“千言万语抵不过一幅图像”。

然而,图像数字化后的数据量之大是非常惊人的。因此图像的处理受到人们的广泛关注。

1.2 研究的背景及意义

图像作为传递信息的媒体和手段是十分重要的。在多媒体计算机系统、电子出版、视频会议、数字化图书馆等许多领域,数字图像都有着广泛的应用。然而用以表示这些图像的数据量很大,使得存储成本很高,尤其不能满足现今网络环境下进行多媒体通信的需求。在这样的背景下,图像编码理论越来越受到广泛的关注,并且迅速发展起来。

图像编码,即图像压缩,就是在保证一定重构质量的前提下,通过去除图像中的各种冗余,以尽量少的比特数来表征图像信息。传统的图像编码是以香农信息论为出发点,用统计概率模型来描述信源。编码实体是像素或像素块,以消除数据相关冗余为目的,由此产生的 JPEG、MPEG-1、MPEG-2、H.261 及 H.263 等编码国际标准已经获得了很大的成功。

信息技术的发展给图像压缩和传输技术带来了更多的机遇和挑战,传统的基于块 DCT 变换压缩编码的图像压缩方法具有以下缺点:①当压缩比较高时在重建图像中会出现方块效应;②压缩比由量化参数控制,调整不方便;③ 压缩的效率不高。而小波变换是在整幅图像中进行,将图像逐级分解到不同分辨率的尺度空间,再对分解系数进行扫描编码,从根本上消除了方块效应。由于扫描编码输出结果位流(bit stream)对重建图像的贡献严格按从重到轻排列,可以通过控制编码输出位的数目来无级控制图像压缩比。另外,基于小波变换的图像压缩方法在重建图像质量接近的情况下压缩倍数比 DCT 变换编码方法高许多。因此近来该方

法在图像压缩中日趋得到广泛应用。在静止图像压缩编码中，DCT 变换已经被小波变换取代，在视频压缩领域，小波变换因为其优良的处理特性也正逐步取代 DCT 变换。

小波分析^{[1][2]}是傅立叶分析发展史上的一个里程碑，被誉为“数学显微镜”。作为一种多分辨率分析方法，小波分析具有很好的时频局部化特性，特别适合按照人眼视觉特性设计图像编码方法，也非常有利于图像信号的渐进传输。在实际应用中，基于小波变换的图像编码，在压缩比和编码质量方面均优于传统的 DCT 变换编码。提升小波变换的提出，为小波图像编码又注入了新的活力，与传统的小波变换相比，它不依赖于傅立叶变换，计算简单，时间、空间复杂度低，易于实现，被称为第二代小波变换。而且，随着各种高效的小波系数量化方法的提出，基于小波的图像编码方法的发展越来越迅速，并取得了非常好的效果。所以，在新的国际编码标准 JPEG2000 和 MPEG-4 中都采用了基于小波变换的图像编码方法。

1.3 小波图像压缩编码研究现状

小波变换是一种全局变换，在时域和频域同时有良好的定位能力，它对图像的高频分量采用逐步精细的时域和空域步长，可以聚焦到被处理图像的任何细节。由于小波变换的这些优点，自二十世纪八十年代开始被广泛应用于图像编码领域，并已成为图像压缩领域的一个研究热点。

一般而言，小波图像编/解码具有如图 1.1 所示的统一结构框架。

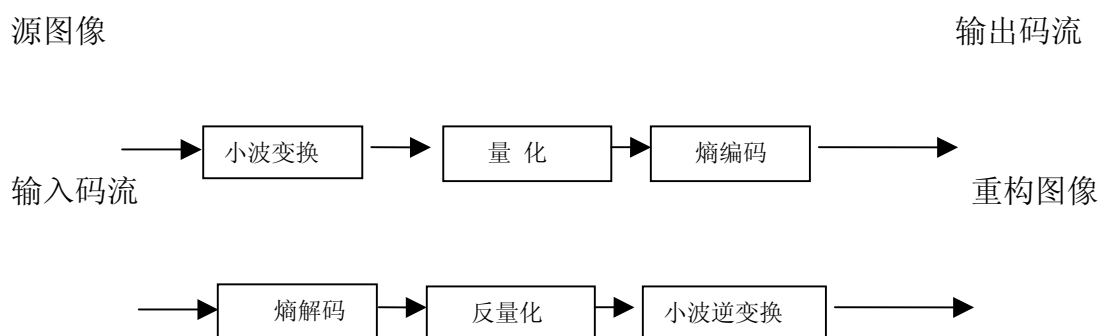


图 1.1 图像编/解码框图

Fig 1.1 graph of image encoding/decoding based on wavelet transform

图像经小波变换后，并没有实现压缩，只是对整幅图像的能量进行了重新分配。事实上，变换后的图像具有更宽的范围，但是宽范围的大数据被集中在一个小区域内，而在很大的区域中数据的动态范围很小。小波变换编码就是在小波变换的基础上，利用小波变换的这些特性，采用适当的方法组织变换后的小波系数，实

现图像的高效压缩的。

熵编码主要有游程编码、霍夫曼编码和算术编码等。而量化为小波编码的核心,其目的是为了为了更好的组织图像经小波变换后的系数,以实现高效压缩。

目前,基于小波变换的图像编码方法的研究热点有以下几个方面:

(1)小波基的选择:选择最优小波基用于图像编码是一个非常棘手的问题,迄今为止尚未发现统一的选择标准。但还是存在一些可参考原则,如平滑性、逼近精度、支撑大小和滤波频率选择等。文献[3][4]已指出:正交小波基的平滑性与消失矩(vanishing moments)对图像压缩效果有一定影响,并且平滑性的影响要高于消失矩。文献[5]验证了滤波器长度小于36的双正交小波基对图像编码的影响,由于滤波器长度决定了计算量的大小和边界扩展所引起的失真度大小,应选择适中的滤波器长度。文献[6]给出了用提升算法构造双正交小波的方法。文献[7]介绍了双正交小波在图像压缩中的应用。

(2)DWT(Discrete Wavelet Transform)的快速算法:自Mallat算法^{[1][5]}被提出后,许多DWT的改进算法相继被提出。Rioul等人^[8]提出了利用FFT进行卷积运算,对长度较大的DWT计算效果较好。Sweldens提出的提升方法^{[9][10]}可以使DWT的计算复杂度在Mallat算法的基础上降低一半,并成为了主流的DWT计算方法。文献[6][7]都对提升算法做了详细介绍。而对于二维DWT不可分离的计算方法,相关研究很少,[11]提出的SCLA算法在这方面取得了很好的成果。

(3)整数小波变换:在图像编码中,图像的像素值是整数,对其实施整数到整数的变换,可以保证信息的无损表示,这一点是现有的浮点离散变换,包括DCT、DFT以及传统的小波变换都不能做到的,因为目各类离散变换,其基函数均为三角函数,因此必然将整数变换到浮点数,而传统的小波变换是基于卷积运算的,其计算量大,对存储空间要求高,变换后的小波系数也是浮点数,受到计算机的表达精度的限制,不能实现图像完全的无损编码,而且硬件上也不易实现。Sweldens提出了一种新的小波构造方法—提升方法^{[9][10]},对该方法稍加改进,就可以很容易得到整数到整数的小波变换^[12]。提升方法的小波变换不依赖于傅立叶变换,计算量小,易于硬件实现,可以实现图像的完全无损编码。文献[13][14]证实了通过构造从整数到整数的小波可以实现真正意义上的图像无损压缩。文献[15]对于选择何种整数小波变换以提高编码效率进行了分析,文献[16]则将整数小波变换中的截断取整看作向信号中引入加性噪声的过程,从修改整数小波变换的角度改善了压缩效果。(4)小波系数量化方法:现有的量化方法主要集中在标量量化和矢量量化两种方法。标量量化方法中,Shapiro提出的嵌入式零树小波EZW算法^[17]是非常有效的方法,它有效地利用了小波系数的特性,实现了图像的可分级编码。之后,在对EZW改进的基础上,又有学者提出了一些新的性能更好的算法,如多级树集合分裂算法SPIHT^[18]、集合分裂嵌入块编码SPECK^[19]等,都取得了很好

的效果。文献[20]对 SPIHT 进行了改进并用于感兴趣区域图像数字水印认证方法。

基于小波变换图像压缩编码如何和人眼的视觉特性相结合,如何选择适合人眼视觉特性的小波基和相应的滤波器组,如何优化量化和编码过程的效率,以及如何处理高、低频图像数据是值得进一步研究的方向。

总之,基于小波的图像压缩编码算法在数字图像处理、信号和信息处理、多媒体传输处理、通信与信息处理以及计算机相关专业技术人员的共同努力下,已取得了丰富的研究成果,同时,也扩展了这一领域的研究内容,具有很大的发展空间,且有越来越多的研究生投入到基于小波的图像压缩算法研究工作中,并发表了大量的文章,如文献[21][22][23][24]

1.4 本文主要内容及章节安排

本文的工作主要是在小波变换基本理论的基础上,研究基于提升方法的整数小波变换,并设计了一套基于整数小波变换的图像压缩编码方案;研究基于零树思想的小波系数量化方法,分析比较现有方法的优缺点,在研究图像经整数小波变换后系数分布规律的基础上,提出了一种改进的零树编码算法---不完全扫描法,对小波系数进行编码,实验结果证明提高编码效率、减少了编码时间、并节约了内存开销。

本文的章节安排如下:

1 绪论,主要介绍课题的背景、意义以及图像编码理论的基本原理和发展现状。

2 图像压缩编码理论概述。主要介绍图像编码的基本原理,图像压缩编码技术的发展,图像编码的基本方法,图像编码性能的评价标准,以及目前比较成功的图像编码国际标准。

3 小波分析的基本理论。介绍了小波发展的历史背景、小波变换的基本理论;详细阐述了多分辨率分析和 Mallat 算法;分析研究了基于提升方法的整数小波变换方法,并对提升算法和传统小波变换进行了比较。

4 基于嵌入式思想的小波系数编码。主要介绍小波变换图像的系数分布特性,阐述基于嵌入式思想的小波系数量化方法: EZW、SPIHT、SPECK。

5 基于整数小波变换的图像压缩编码方案。介绍了基于提升算法的整数小波的选取,提出了改进的零树编码方法—不完全扫描法。

6 实验结果及分析,对算法进行仿真实验,并对实验结果进行了分析。

7 全文总结及展望,对本文的工作进行了总结,并指出不足之处和今后的研究方向。

2 图像压缩编码理论概述

图像可以定义为景物在某个介质上的再现,例如图片、电影、传真、电视等介质都可以使人们获得图像信息。把图像信息传送到远方或是存储图像信息的过程,统称为图像传输。图像分为静止图像和视频图像(或活动图像)两大类。属于静止图像传输的通信业务有传真、电子邮件、智能用户电报、图文电视、可视图文等。这类图像传输的速率比较低,大部分可以在普通电话线或一般低速的信道上传输。属于视频图像传输通信业务的有可视电话、会议电视、电视传送等。这类图像传输要求高速度、宽频带的信道。要实时地传送活动的图像,在屏幕上显示出来,必须要用大容量的信道,如微波、同轴电缆、卫星、光缆等。

2.1 图像压缩编码的基本原理

从信息论观点来看,图像作为一个信源,描述信源的数据是信息量(信源熵)和信息冗余量之和。图像压缩编码的原理是尽量去除图像数据中的冗余,从而实现对图像数据的压缩编码。目前实现图像压缩的途径有:

空间冗余(Spatial Redundancy):一幅图像中,景物一般不会突变,因此相邻像素之间有很大的相似性,称为空间冗余。可以用 DPCM, 抽样与内插, DCT 变换等方法去除。

时间冗余(Temporal Redundancy):序列图像中相邻帧之间有很强的相关性,这反映为时间冗余。可以用帧间预测,运动补偿,帧间抽样内插等方法去除冗余。

信息熵冗余:信息熵是指数据所携带的信息量,高于信息熵的数据量是信息熵冗余,可采用 Huffman 编码、算术编码等去除冗余。

视觉冗余:人类的视觉系统对于图像场的注意是非均匀和非线性的,对变化平缓区域的误差较变化剧烈区域敏感,即对低频成分敏感,则高频成分中存在冗余。常采用低频精细量化的方法去除冗余,即把人眼不敏感的高频信息丢弃。

其他冗余:如结构冗余、知识冗余等。

图像压缩编码解决的主要问题是采用一种新的表达方式来缩小表示图像所需要的数据量。一般的图像编码算法是在对数字图像进行大量统计分析的基础上,充分利用图像自身相关性强的特点,减少图像信息的冗余度,降低数码率。

2.2 图像压缩编码技术的发展

图像压缩编码技术可以追溯到 1948 年提出的电视信号数字化,到今天已经有半个多世纪的历史了。压缩编码算法繁多,按照复原图像是否与原图包含的信息

量一致为标准分类, 图像压缩编码可分为两大类。第一类压缩过程是可逆的, 从压缩后的图像能够完全恢复出原来的图像, 在压缩和解压缩过程没有信息损失, 称为无损压缩或无失真压缩。第二类压缩过程是不可逆的, 无法恢复出原图像, 信息有一定的丢失, 称为有损压缩或限失真压缩。

总体来说压缩编码技术的发展可以分为两个阶段: 经典编码方法和现代编码方法^[25]。

经典编码方法始于 Shannon 的信息论。1948 年 Shannon 在《通信的数学原理》中提出信息率——失真函数概念, 1958 年又进一步确立了失真率理论, 从而奠定了信源压缩编码的理论基础。经典编码方法是建立在仙农信息论基础之上的, 是以经典集合论为工具, 用概率统计模型来描述信源, 其压缩思想基于数据统计, 因此只能去除数据冗余, 属于低层压缩编码的范畴。该阶段代表方法有: DPCM、PCM、Huffman 编码、算术编码、游程编码、预测编码、离散余弦变换 DCT (Discrete Wavelet Transform) 变换编码、混合编码等。第二个阶段是以小波变换、分形几何理论、神经网络等技术为基础的新颖高效的现代编码方法阶段; 该阶段编码方法突破以仙农信息论为基础的框架的束缚, 充分考虑了人眼视觉特性及信源特性, 通过去除内容冗余来实现数据压缩。该阶段编码方法以小波变换编码研究成果较为突出, 已被 JPEG2000、MPEG-4 等国际图像编码标准所采纳, 成为当前图像压缩的研究热点。

图像压缩编码技术研究工作总体上主要朝两个方向发展:

一是更好地实现现有的图像编码国际标准建议

(1) 研制出集成度更高、性能更好的图像编码专用芯片 (ASIC), 使编码系统成本更低, 可靠性更高。

(2) 解决好现有的图像编码系统开发中的技术问题。例如: 提高图像质量, 提高抗误码性能、降低编解码算法复杂度等。

二是对图像编码理论和图像编码方法的研究

如预测编码与分块 DCT 相结合的经典混合编码方法既不能有效地反映图像的固有统计特性, 也不能有效地反映人们的视觉特性。

基于小波变换的图像编码、模型编码、利用人工神经网络的压缩编码、利用分形几何的图像编码等新颖高效的现代编码方法的研究还有待人们的进一步深入, 许多关键技术还需要有新的突破。

2.3 图像编码的基本方法

图像压缩编码的方法有许多种, 按照不同的标准可以分成不同的类。按信息是否有损失可以分为两大类: 无损压缩编码和有损压缩编码。无损压缩编码符合

理想要求, 但压缩比都不大。如哈夫曼编码、算术编码、游程编码等。有损压缩编码(也称不可逆压缩编码)编码后图像信息有损失, 且压缩比越高, 图像失真越明显。如预测编码, 小波变换编码, DCT 变换编码, 分形编码, 神经网络编码等。

2.3.1 熵编码

常用的熵编码有霍夫曼编码(Huffman coding)、算术编码(arithmetic coding)、游程编码(RLC, run-length coding)。^{[18][19]}

霍夫曼编码是可变字长编码(VLC)的一种。Huffman 于 1952 年提出一种编码方法, 该方法完全依据信源字符出现概率来构造平均码子长度最短的异字头码字, 有时称之为最佳编码, 一般就叫做 Huffman 编码。它将最短的编码赋给那些最频繁出现的字符, 而概率小的字符则分配较长的码字, 从而提高编码效率。Huffman 编码是图像压缩中最重要的一种编码方法之一。其编码步骤如下:

(1) 将输入符号按出现的概率由大到小(顺序排列相同概率的符号可以任意颠倒排列位置)。

(2) 将最小的两个概率相加得到一个新的概率, 再去掉那两个最小的概率, 形成一个新的概率集合。再按第一步的方法重排。如此重复直到两个概率为止。

(3) 分配码字。码字分配从最后一步开始反向进行, 对最后两个概率, 一个赋予“0”码字, 一个赋予“1”码字。

该编码结构实际上是一棵二叉树, 码字都是从根出发排列的。概率大的符号分配较短的码字, 概率较小的符号分配较长的码字, 从而提高其编码效率。

算术编码的概念是 Elias 在六十年代初期提出来的, Rissanen 和 Pasco 首次介绍了它的实用技术。到八十年代才出现了算术编码的具体实现方法。算术编码是另一种利用信源编码概率分布特性, 能够趋近熵极限的编码方法。尽管它也是对出现概率大的符号采用短码, 对出现概率小的符号采用长码, 但其编码原理与 Huffman 编码却不相同, 而且在信源概率分布比较均匀的情况下其编码效率要高于 Huffman 编码。它和 Huffman 编码的最大区别在于它不是使用整数码。

算术编码的方法是将被编码的每一个信源消息按其概率大小表示成实数轴 $[0, 1]$ 之间的一个区域(或称之区间), 概率越大, 所占间隔越大, 表示这一间隔所需的二进制分数值位数就越小, 反之就越大。算术编码所产生的码字实际上就是一个二进制分数值的指针, 该指针指向所编符号对应的概率区间, 该区间为半开区间, 包括左端点, 不包括右端点。算术编码的规则如下:

(1) 在初始状态, 编码指针 $C=0$, 子区间长度 $L=1.0$

(2) 在编码状态, 编码指针 $C=\text{原始码点 } C + \text{原子区间长度 } L \times \sum_k P_i$

子区间长度 L = 原子区间长度 $L \times P_i$ 。

其中 P_i 和 $\sum_k P_i$ 分别为所编符号对应的概率和累计概率。

(3) 取最后一个被编码符号编码指针的二进制分数值作为对整个信源符号编码集的编码结果。

游程编码(Run-length Coding)是一种相对简单的编码技术, 由于图像中各像素之间一般都存在相关性, 特别是由计算机生成的图像和大部分二值图像, 它们往往在某些区域具有相同的像素值。主要思想是将一个相同值的连续串用一个代表串长和值的组合码字(Run, Level)来代替。其中游程(Run)表示系数系列中连零系数的长度;量化电平(Level)表示连零系数之后第一个非零系数的值。

H.261 建议中使用的游程编码就是上面所述的二维游程编码; H.263 建议对此作了改进, 使用的是三维游程编码(Last, Run, Level), 其中的 Last 表示此码字组合是否是本块中最后一个非零系数(“0”表示不是, “1”表示是), Run 和 Level 含义和上面所说的相同。

游程编码一般不直接应用于多灰度图像编码, 但比较适合于二值图像的编码, 如传真图像的编码等。为达到较好的压缩效果, 游程编码经常和其他一些编码方法混合使用。H.263 中游程编码和 DCT 及 Huffman 方法一起使用, 对分块作完 DCT 及量化后的频域图像数据作“Z”型扫描, 然后作游程编码, 对游程编码的结果再做 Huffman 编码

2.3.2 变换编码

变换编码(transform coding)是通过信号变换来消除图像数据空间相关性的一种有效方法^[28]。常用的有 DCT 变换和小波变换。

离散余弦变换编码 DCT(Discrete Cosine Transform)是先将整体图像分成 $N \times N$ 像素块, 然后对 $N \times N$ 像素块逐一进行 DCT 变换。由于大多数图像的高频分量较小, 相应于图像高频成分的系数经常为零, 加上人眼对高频成分的失真不太敏感, 所以可用更粗的量化, 因此传送变换系数所用的数码率要大大小于传送图像像素所用的数码率。到达接收端后再通过反离散余弦变换回到样值, 虽然会有一定的失真, 但人眼是可以接受的。

小波变换编码 DWT(Discrete Wavelet Transform)是先对图像作小波变换, 然后对变换后得到的系数进行编码。由于小波能有效去除像素间的空间冗余, 故能对图像进行有效压缩。现在, 小波变换编码已成为图像处理领域研究的热点。

2.3.3 预测编码

预测编码可以在一幅图像内进行, 也可以在多幅图像之间进行。预测编码实际上是基于图像数据的空间和时间冗余特性, 用相邻的已知像素(或图像块)来

预测当前像素（或图像块）的值，然后再对预测误差进行量化和编码。预测编码的关键在于预测算法的选取，这与图像信号的概率分布有很大关系。实际中常根据大量的统计结果来设计最佳的预测器，有时还使用自适应预测器。帧内预测编码所使用的 DPCM（差分脉冲编码调制），其优点是算法简单，易于硬件实现，缺点是对信道噪声及误码敏感，同时压缩效率也不高。帧间预测编码主要利用视频序列相邻帧间的相关性，即图像数据的时间冗余来达到压缩的目的，通过运动补偿编码来消除两帧间的冗余，能获得很高的压缩倍数。对于视频图像而言，由于相邻帧之间的时间间隔很短，根据统计实验的结果，大部分像素保持不变，或只有微小的运动。因此可用前一帧相应位置的像素加上微小的运动矢量对当前像素进行预测，这种方法的关键问题是要估算出运动矢量，运动估计的方法有块匹配、像素递归法、模型法等。

2.3.4 运动补偿编码

运动补偿编码是视频压缩的关键技术之一。目前已成为国际标准的视频压缩方法中，从 H.261, H.263, MPEG-1 到 MPEG-2, 都无一例外地使用了运动补偿（简称为 MC）技术。运动补偿技术的主要内容包括：

- （1）图像分割。将视频图像分割成静止部分和运动部分。
- （2）运动检测与估值。检测运动类型（平移、旋转和缩放等），估计运动物体的位移值。
- （3）运动补偿。用位移估值进行运动补偿预测。
- （4）预测信息编码。对预测信息（例如，运动矢量）进行编码，作为边信息传送给接收端。

图像分割是运动补偿预测的基础，但实际上物体的形状多种多样，要把图像中不同运动的物体分割出来是比较困难的。通常采用两种比较简单的方法。一是把图像划分成矩形子块，适当选择子块大小，把子块分为静止和运动两类，估计出运动子块的位移，进行预测编码。这种方法称之为块匹配法（BMA-block match algorithm）。另一种方法是对每个像素的位移都进行递归估计，因此称之为像素递归法（PRA-pixel recursion algorithm）。这两种算法各有其特点，通常像素递归法较块匹配法估计精度高，对多运动画面的适应能力也强，但只能跟踪较小的位移（2-3 像素/帧），且实现复杂。块匹配法虽然估计精度低于像素递归法，但由于其位移跟踪能力强（不低于 6-7 像素/帧），且实现简单，因此在实际中获得了广泛的应用。

2.4 图像编码方法的评价标准

对于一个图像编码方法的评价，主要包括以下两个方面：一是对重构图像质量的评价，二是对图像编码效率的评价。

通常的图像编码都是失真编码,这就需要有一个对重构图像质量进行评价的标准,希望在传输码率尽量低的前提下能尽量保证重构的图像具有较高的图像质量,这对图像编码方法的发展具有很好的指导意义。图像质量的含义主要包含了两层意义:一是重构图像的失真度,即重构图像与原图像的偏离程度;二是图像的可读性,即人们能从图像中获得的信息。一般情况下,人眼是图像信息的接收者,但由于目前对人的视觉系统的性质的理解还很有限,还不能对图像的失真度和可读性做定量描述,所以在图像质量评价方法上,除了主观的评价方法外,还需要客观的评价方法。

主观评价方法是由评价者直接对一幅图像进行观察,从感觉上去度量其失真程度,给出质量评价分数,对所有评价者给出的分数进行加权平均,所得结果即为主观评价结果,有两种尺度:绝对尺度和相对尺度。绝对尺度可以分为 5 个等级:很好,较好,一般,较差,很差;相对尺度也分为 5 个等级:最好的,高于平均水平,平均水平,低于平均水平,最差的。这种评价结果必然符合人的视觉感受,相对于客观评价方法是可靠的,但它使用起来不方便,一方面不能用数学模型对其进行描述,不能直接用于图像编码过程中的质量评价和控制,另一方面,主观评价容易受到评价者的主观因素影响,如年龄、教育程度和性格等,所以大多数情况下还是以客观评价方法对图像质量进行评价。

所谓客观评价方法,就是定义一个数学公式,然后对待评价的图像进行运算,得到一个唯一的数字量作为测度结果,通常用于评价图像的失真度。

图像压缩的好坏通常由以下几个量表征:

(1) 压缩比 CR(Compression Ratio)。

$$CR = \text{压缩前图像所含比特数} / \text{压缩后图像所含比特数} \quad (2.1)$$

图像的压缩效率通常用压缩比来衡量,压缩比越高,图像压缩越大,反之亦然。

(2) 编码及解码时间。

编解码的时间越短越好,但要实现比较好的压缩效果,必然要采用较复杂的算法,运算量的增加必然使编解码时间增加,因此这是一对矛盾,在实际应用中要综合考虑。同时,编码和解码也存在一定的矛盾,一般由于编码端环境的限制,对编码的时间要求更严格,也常采用尽量减少编码端的计算,将这些计算放到解码端进行处理。

(3) 重建图像质量。

人眼是重建图像主观评价的工具,通常至少选 20 人作为评判者,以保证结果的可靠性;在实际中常用一些客观参数作为评判重建图像质量的一个度量。这些参数有 PSNR, MSE。

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=0}^M \sum_{j=0}^N [\psi(i, j) - \psi'(i, j)]^2 \quad (2.2)$$

其中 $M \times N$ 是图像的尺寸, $\psi(i, j), \psi'(i, j)$ 分别表示原图像和重建图像在 (i, j) 处的灰度值。

$$\text{PSNR} = 10 \log_{10}(l^2 / \text{MSE}) \quad (2.3)$$

其中 l 表示图像灰度值的量化级数, 若 8 比特量化的话, 则 $l=255$ 。

应该注意的是, MSE 和 PSNR 是从总体上反映原始图像和重构图像的差别, 并不能反映局部。有时候同样的信噪比, 视觉效果还是会有一定的差异, 这主要是由于误差的均匀程度造成的。一般来讲, 误差均匀时视觉效果好, 反之视觉效果不好。大多数情况下, 都可以用 PSNR 对图像质量进行评价, 但有时其结果可能与主观评价结果不相符。

2.5 图像编码的国际标准

近十几年来, 图像编码技术得到了迅速发展和广泛应用, 并且日臻成熟, 其标志就是几个关于图像编码的国际标准的制定, 即国际标准化组织 ISO 和国际电工委员会 IEC 关于静止图像的编码标准 JPEG 和 JPEG2000、关于活动图像编码标准 MPEG-x 系列和国际电信联盟 ITU-T 关于可视电话、电视的视频编码标准 H.26x 系列。这些图像编码标准融合了各种性能优良的图像编码方法, 代表了目前图像编码的发展水平。

1、JPEG

JPEG(Joint Photographic Expert Group) ^[29] 是 ISO/IEC 联合图像专家组制定的静止图像编码标准, 是适用于连续色调(包括灰度和彩色)静止图像压缩的国际标准。JPEG 算法共有四种运行模式, 其中一种是基于空间预测(DPCM)的无损压缩算法, 另外三种是基于 DCT 的有损压缩算法。

- (1) 无损压缩算法, 可以保证无失真的重构原始图像。
- (2) 基于 DCT 的顺序模式, 按从上到下, 从左到右的顺序对图像进行编码, 称为基本系统。
- (3) 基于 DCT 的渐进模式, 指对一幅图像按由无损到有损进行编码。
- (4) 分层模式, 以各种分辨率对图像进行编码, 可以根据不同的要求, 获得不同分辨率的图像。

这些模式中的基本系统(顺序模式)是各个 JPEG 都必须支持的模式, 它提供了适合大多数应用场合的简单高效的图像编码方案。增强系统(递进模式、分层模式)是基本系统的扩充或增强, 在增强系统中必须包括基本系统。

灰度图像中只有一个亮度分量, 而彩色图像有一个亮度分量和两个色度分量,

对彩色图像编码时可以按照对灰度图像的编码方法对每一个分量进行编码。

2、JPEG2000

与以往的 JPEG 标准相比, JPEG2000^{[30][31][32]}有了一个很大的飞跃,它有许多原来的标准所不可比拟的优点。JPEG2000 与传统的 JPEG 最大的不同,在于它放弃了 JPEG 所采用的以 DCT 变换为主的分块编码方式,而改为以小波变换为主的多分辨率编码方式。

首先, JPEG2000 能实现无损压缩(lossless compression)。在实际应用中,有一些重要的图像,如卫星遥感图像、医学图像、文物照片等,通常需要进行无损压缩。对图像进行无损编码的经典方法---预测法已经发展成熟,并作为一个标准写入了 JPEG2000 中。

JPEG2000 还有一个很好的优点就是误码鲁棒性(robustness to bit error)好。因此使用 JPEG2000 的系统稳定性好,运行平稳,抗干扰性好,易于操作。JPEG2000 能实现渐进传输(progressive transmission),这是 JPEG2000 的一个极其重要的特征。它可以先传输图像的轮廓,然后逐步传输数据,不断提高图像质量,以满足用户的需要,这在网络传输中具有非常重大的意义。使用 JPEG2000 下载一个图片,用户可先看到这个图片的轮廓或缩影,然后再决定是否下载它。而且,下载它时可以根据用户需要和带宽来决定下载图像质量的好坏,从而控制数据量的大小。

JPEG2000 的另一个极其重要的优点就是感兴趣区域(ROI, Region Of Interest)特性。用户在处理的图像中可以指定感兴趣区域,对这些区域进行压缩时可以指定特定的压缩质量,或在重构时指定特定的解压缩要求,这给人们带来了极大的方便。在有些情况下,图像中只有一小块区域对用户是有用的,对这些区域采用低压缩比,而感兴趣区之外采用高压缩比。在保证不丢失重要信息的同时,又能有效地压缩数据量,这就是基于感兴趣区域的编码方案所采取的压缩策略。基于感兴趣区域的压缩方法的优点,在于它结合了接收方对压缩的主观要求,实现了交互式压缩。

JPEG2000 标准还充分考虑了人眼视觉特性,增加了视觉权重和掩膜,这样在不损害视觉效果的情况下,可以大大提高压缩效率。

3、H.261 标准

H.261^[33]是 ITU-T 针对可视电话和会议电视、窄带 ISDN 等要求实时编解码和低延时应用提出的一个编码标准。该标准所包含的比特率为 $p \times 64$ kbps,其中 p 是一个整数,取值范围为 1-30。H.261 要求输入的视频格式为 CIF 或者 QCIF。H.261 信源编码算法采用 DPCM/DCT 混合编码算法, H.261 建议在编码方面主要采用了三种技术来实现压缩:预测编码,变换域编码,游程熵编码。

4、H.263 标准^[34]

1995 年 ITU-T 针对低比特率视频应用制定了 H.263 标准。随后几年, ITU-T 又对其进行了多次补充, 以提高编码效率, 增强编码功能。补充修订的版本有 1998 年的 H.263+^[35], 2000 年的 H.263++^[33]。H.263 系列标准特别适合于 PSTN 网络、无线网络与因特网等环境下的视频传输。其标准的输入图像格式可以是 S-QCIF、QCIF、CIF、4CIF 或者 16CIF 的彩色 4:2:0 亚取样图像。同 H.261 一样, H.263 采用了运动补偿和 DCT 编码方法, 但它参照 MPEG 标准引入以 I 帧、P 帧、PB 帧(选项)三种帧模式和 INTER(帧间编码)、INTRA(帧内编码)两种编码模式。H.263 在 H.261 基本编码算法的基础上又提供了四种可选编码模式, 以进一步提高编码效率。

5、MPEG-1^[36]

MPEG-1 标准于 1993 年 8 月公布, 全称为“用于数字存储媒体的 1.5Mbps 以下活动图像及其相关音频编码”(Coding of Moving image and Associated Audio for Digital Storage Media at up to 1.5Mbps), 是 MPEG 推出的第一个标准, 主要应用于视频信号(如电影节目)的数字化压缩存储, 其图像质量与家用录像机相近。MPEG-1 的原理框图与 H.261 有许多共同之处, 但有重要改进。由于有运动补偿的帧间内插可以获得更好的压缩效果, 以及随机查询对于数字存储媒体在电视图像重放以及节目编辑中极为重要, MPEG-1 将数字图像帧分为三类: I(Intra)图像帧仅作帧内编码; P(Prediction)图像帧以过去的参考图像帧进行正方向的帧间预测编码; B(Bidirection prediction)图像帧则根据过去的以及后来的参考图像进行时间域正负双方向进行帧间预测编码。

6、MPEG-2^[37]

MPEG-2 是 1994 年推出的, 全称为“活动图像及伴音的通用编码”(Generic coding of Moving Pictures and Associated Audio), 同 MPEG-1 标准一样, MPEG-2 标准也包括系统、视频和音频等部分内容; 它克服并解决了 MPEG-1 不能满足日益增长的多媒体技术、数字电视技术对多分辨率和传输率等方面的技术要求的缺陷。与 MPEG-1 相比功能上有了很大的补充, 图像质量上也有了很大的提高, 它不仅支持普通的 CIF、CCIR-601 等分辨率格式, 还可支持高清晰度分辨率; 不仅支持面向存储媒介的应用, 还广泛支持各种通信环境下数字视频信号的编码与传输, 如卫星广播、数字地面广播等; 不仅支持恒定比特率传输模式, 还支持变比特率传输模式。MPEG-2 另一个重要特色是其比特流的可分级性, 这意味着解码器可以忽略比特流中的增强部分, 只解码全部比特流中的基本部分, 仍可得到有用的图像序列, 只是分辨率要低一些。

7、MPEG-4^[38]

MPEG-4 公布于 1998 年, MPEG-4 与 MPEG-1 和 MPEG-2 有很大的不同。

MPEG-4 不只是具体压缩算法，它是针对数字电视、交互式绘图应用、交互式多媒体等整合及压缩技术的需求而制定的国际标准。MPEG-4 的编码理念是：MPEG-4 标准同以前标准的最显著的差别在于它是采用基于对象的编码理念，即在编码时将一幅景物分成若干在时间和空间上相互联系的视频音频对象，分别编码后，再经过复用传输到接收端，然后再对不同的对象分别解码，从而组合成所需要的视频和音频。这样既方便我们对不同的对象采用不同的编码方法和表示方法，又有利于不同数据类型间的融合，并且这样也可以方便的实现对于各种对象的操作及编辑。例如，我们可以将一个卡通人物放在真实的场景中，或者将真人置于一个虚拟的演播室里，还可以在互联网上方便的实现交互，根据自己的需要有选择的组合各种视频音频以及图形文本对象。

3 小波分析理论

3.1 引言

小波变换是一种新的分析方法,它是继傅立叶变换之后纯粹数学(基础数学)和应用数学完美结合的又一光辉典范。小波变换的产生、发展、完善和应用始终受惠于计算机科学、信号处理、图像处理、应用数学、纯粹数学、物理学、地球科学等众多科学领域和工程技术应用领域的专家、学者和工程师们的共同努力。正因为如此,小波变换现在已成为科学研究和工程技术应用中涉及面极其广泛的一个热门话题。当然,这也决定了大家对小波变换的不同看法。数学家说,小波理论是函数空间的一种优美的表示理论;信号处理专家认为,小波变换是非平稳信号时-频(Time-Frequency Analysis)分析的新理论;图像分析和处理领域的专家又认为,小波变换是数字图像处理的空间-尺度分析(Space-Scale Analysis)和多分辨率分析(Multi-Resolution Analysis)的有效工具;地球科学和故障诊断的学者认为,小波变换是信号奇性识别之位置-尺度分析(Position-Scale Analysis)的一种新技术;微局部分析家觉得小波变换体现了微局部分析方法中时间-尺度分析(Time-Scale Analysis)的新思想。

3.2 小波发展的历史背景

小波分析的历史是数学家和工程师共同创造的,它的产生、发展以及成熟凝结了许多人的智慧和劳动。这种分析方法的提出,可以追溯到1910年,Haar提出的规范正交基。1936年Little wood和Paley对Fourier级数建立的L-P理论。其后,Calderon于1975年给出了抛物型空间上的原子分解,它的离散形式已接近小波展开,只是无法得到组成正交基的结论。1981年Stromberg对Haar系数进行了改进,证明了小波函数的存在性。1982年Battle在构造量子场论时使用了类似Calderon再生公式的展开,值得注意的是,1984年法国地球物理学家Morlet在分析地震波的局部性质时,发现传统的Fourier变换难以达到要求,因此,他把小波概念引入了信号分析中对信号进行了分解。随后,理论物理学家Grossman对Morlet的这种信号按一个确定的函数的伸缩、平移展开的可能性进行了研究,为小波分析开了先河。

真正的小波热开始于1986年,当时Meyer创造性地构造出具有一定衰减特性的光滑函数 Ψ ,它的二进制伸缩平移构成 $L^2(\mathbb{R})$ 的规范正交基^[39]。在那之前,人们认为具有如此良好性质的小波函数是一个数学神话甚而对其存在的真实性发生了怀疑。1989年,Meyer出版的《小波算子》是小波理论这一新兴学科产生的

标志, 该书也是当前最权威、最系统的小波分析理论的专著。继 Meyer 提出小波变换之后, Lemarie 和 Battle 又分别绘出了具有指数衰减的小波函数。

1987 年 Mallet 巧妙地将计算机视觉领域内多分辨率分析的思想引入到小波函数的构造中, 从而成功地统一了在此之前的 Stromberg, Meyer, Lemarie 和 Battle 提出的具体小波函数的构造。1988 年, Daubechies 构造了紧支正交小波基, 从此小波分析的理论初步形成。Coifman 和 Wickerhar 在深化 Mallet 算法的基础上提出小波包理论, 是信号处理领域又一重大发展, 他提出了最优基的选择, 同样如小波变换一样存在快速有效的小波包算法。1992 年, Mallat 将小波分析用于边缘检测, 取得了较好的效果。1993 年, Swelden 等提出了第二代小波基的构造方法: Lifting scheme 算法。1993 年提出了多小波, 从而将小波分析推向了一个新的高潮。

在信号处理领域, Veneli 等系统地研究了小波和滤波器组的关系, 提出了两通道完全重建滤波器组实现离散小波变换和 Mallet 算法以及构造正交、双正交小波基的方法。人们对小波的研究已经从一维扩展到高维。近来又有许多学者致力于多小波的研究。经过这十多年的发展, 小波分析已在理论和方法以及应用上取得了突破性的进展。从信号处理的角度来看, 小波分析优于其他时频分析的地方是它对于时域和频域同时具有良好的局部化的性质, 因此, 成为非平稳信号(如图像信号)处理的强有力的工具。

事实上小波分析的应用领域十分广泛, 它包括: 数学领域的许多学科; 信号分析、图像处理; 量子力学、理论物理; 军事电子对抗与武器的智能化; 计算机分类与识别; 音乐与语言的人工合成; 医学成像与诊断; 地震勘探数据处理; 大型机械的故障诊断等方面; 例如, 在数学方面, 它已用于数值分析、构造快速数值方法、曲线曲面构造、微分方程求解、控制论等。在信号分析方面的滤波、去噪声、压缩、传递、边界的处理与滤波、时频分析、信噪分离与提取弱信号、求分形指数、信号的识别与诊断以及多尺度边缘检测等。在图像处理方面的图像压缩、分类、识别与诊断, 去污等。在医学成像方面的减少 B 超、CT、核磁共振成像的时间, 提高分辨率等。在工程技术等方面的应用。包括计算机视觉、计算机图形学、曲线设计、湍流、远程宇宙的研究与生物医学方面。

小波分析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方法, 小波域纹理模型方法, 小波变换零树编码压缩, 小波变换向量压缩等。

3.3 小波变换的基本理论

3.3.1 傅立叶变换及其存在的不足

傅立叶变换是众多科学领域（特别是信号处理、图像处理、量子物理等）里最重要的工具之一。傅立叶分析是指傅立叶变换和傅立叶级数。

傅立叶变换定义为：

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (3.1)$$

傅立叶逆变换定义为：

$$f(t) = 1/2\pi \int_{-\infty}^{\infty} e^{j\omega t} F(\omega) d\omega \quad (3.2)$$

在实际应用中为了在计算机上实现信号的频谱分析及其他方面的处理，对信号的要求是：在时域和频域是离散的，且都是有限长。这就要作离散傅立叶变换（DFT）。离散傅立叶变换的定义为：

$$X(k) = F(f_n) = \sum_{n=0}^{N-1} f_n e^{-j\frac{2\pi k}{N}n} \quad k=0,1,\dots,N-1 \quad (3.3)$$

离散傅立叶逆变换的定义为：

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi k}{N}n} \quad n=0,1,\dots,N-1 \quad (3.4)$$

傅立叶分析不能反映随时间变化的频率，Gabor 提出了加窗傅立叶变换(STFT)的概念。

傅立叶变换是时域到频域相互转化的工具，从物理意义上讲，傅立叶变换的实质是把 $f(t)$ 这个波形分解成许多不同频率的正弦波的叠加和。这样我们就把对 $f(t)$ 的研究转化为对其权系数（即 $F(\omega)$ ）的研究。

虽然傅立叶变换能够将信号的时域特征和频域特征联系起来，能分别从信号的时域和频域观察，但却不能把二者有机地联系起来。这是因为信号的时域波形中不包含任何频域信息。而其傅立叶谱是信号的统计特性，从其表达式中也可以看出，它是整个时域内的积分，没有局部化分析信号的功能，完全不具备时域信息，也就是说，对于傅立叶谱中的某一频率，不知道这个频率是在什么时候产生的。这样在信号分析中就面临一对最基本的矛盾：时域和频域的局部化矛盾。小波变换具有时域局部化和频域局部化的特点，解决了傅立叶变换的不足，因此，小波变换成为信号分析中的重要工具。

3.3.2 小波的概念

小波理论^{[13][19][26][27]}的发展经历了不同的阶段。Grossman 和 Morlet 于 1984 年提出了小波理论和连续小波变换，最先使用了“小波”这个术语并给以严格的数学

定义。所以,一般认为,小波是由 Grossman 和 Morlet 首先定义的。

定义 2.1 小波是 $L^2(R)$ 空间的一个函数 ψ , 由 ψ 构成的序列 $2^{j/2}\psi(2^j x - k)$, $j, k \in Z$ 构成 $L^2(R)$ 空间的一个正交基。

一旦我们确定了这样一个函数 $\psi \in L^2(R)$, 如果它满足容许条件:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \text{ 则称函数 } \psi \text{ 为基本小波(Basic Wavelet)或母小波(Lazy}$$

wavelet)。

设 $\psi \in L^2(R)$ 是基本小波, 对 ψ 进行平移 (Translation) 和伸缩 (Dilation) 变换, 得:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \quad a, b \in R, \quad a \neq 0 \quad (3.5)$$

我们称 $\psi_{a,b}(x)$ 为连续依赖于参数 a 、 b 的小波, 称为分析小波。

在向量空间中, 基是一组线性无关的向量, 向量空间的任何一个向量都可以用这些基向量的线性组合来表示, 即 $v = \sum_k a_k v_k$ 。同时, 一个向量空间可以有很多不同的基, 但是这些基中的向量个数是相同的, 即向量空间的维数。

同理, 对于函数空间 $L^2(R)$, 如果存在函数列 $\{\phi_k(t)\} \subset L^2(R)$ 构成 $L^2(R)$ 的基, 那么 $L^2(R)$ 中的任意函数 $f(t)$ 都可以表示成 $f(t) = \sum_k a_k \phi_k(t)$ 。

3.3.3 连续小波变换

函数 $f \in L^2(R)$ 的连续小波变换或积分小波变换定义为:

$$Wf(a,b) = \langle f(t), \psi_{a,b}(t) \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{|a|}} \psi^*\left(\frac{t-b}{a}\right) dt \quad (3.6)$$

由式 (3.6) 可以看出, 函数 $f(t)$ 的连续小波变换就是函数 $f(t)$ 与小波 $\psi_{a,b}(x)$ 的内积。

在小波变换的定义中, 小波变换信号是两个变量 a 和 b 的函数, 其中 a 是尺度参数, b 是平移参数。小波变换中的尺度参数相当于一个比例尺的作用, 大比例尺反映细节信息, 小比例尺反映的是非细节的整体信息。小波中的平移参数意味着这种变换可以在整个时域中进行。

函数 $f \in L^2(R)$ 的连续小波逆变换定义为:

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle f, \psi_{a,b} \rangle \psi_{a,b}(x) \frac{da db}{a^2} \quad (3.7)$$

$$\text{其中 } C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty。$$

根据小波的容许性条件, 即 $\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty$, 可以得到 $\hat{\psi}(0) = 0$ 。

由于 $\hat{\psi}(\omega) = \int_{-\infty}^{+\infty} \psi(t) e^{-j\omega t} dt$, 令 $\omega = 0$, 则可以得到:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (3.8)$$

从式 (3.8) 可以得到, “小波”意味着该函数的积分为 0, 窗口函数具有有限长度, 即具有紧支性, 同时在时间轴上下波动。

3.3.4 傅立叶变换与小波变换的比较

小波分析是傅立叶分析思想方法的发展与延拓。它自产生以来, 就一直与傅立叶分析密切相关。它的存在性证明, 小波基的构造以及结果分析都依赖于傅立叶分析, 二者是相辅相成的。二者相比较主要有以下不同:

- (1) 傅立叶变换的实质是把能量有限信号 $f(t)$ 分解到以 $\{e^{j\omega t}\}$ 为正交基的空间上去; 小波变换的实质是把能量有限信号 $f(t)$ 分解到 W_j 和 V_j ($j=1,2,\dots,J$) 所构成的空间上去 (W_j 是小波函数空间, V_j 是尺度函数空间, 在下文还有详细介绍)。
- (2) 傅立叶变换用到的基本函数只 $\sin(\omega t)$, $\cos(\omega t)$, $\exp(j\omega t)$, 具有唯一性; 小波分析用到的函数 (即小波函数) 则具有不唯一性, 同一个工程问题用不同的小波函数进行分析有时结果相差甚远。小波函数的选择是实际应用的一个难点问题, 也是小波分析研究中的一个热点问题。
- (3) 在频域中, 傅立叶变换具有较好的局部化能力, 特别是对于那些频域成分比较简单的确定性信号, 傅立叶变换很容易把信号表示成各频率成分的叠加和的形式, 例如, $\sin(\omega_1 t) + 2.3 \sin(\omega_2 t)$ 。但在时域中, 傅立叶变换没有局部化能力, 即无法从 $f(t)$ 的傅立叶变换 $F(\omega)$ 中看出 $f(t)$ 在任一时间点附近的频率状态。事实上, $F(\omega)d\omega$ 是关于频率为 ω 的谐波分量的振幅, 在傅立叶展开式中, 它是由 $f(t)$ 的整体性态所决定的。
- (4) 在小波分析中, 尺度 a 的值越大相当于傅立叶变换中 ω 的值越小。
- (5) 在短时傅立叶变换中, 变换系数 $S(\omega, \tau)$ 主要依赖于信号在 $[\tau-\delta, \tau+\delta]$ 片段中的情况, 时间宽度是 2δ (因为 δ 是由窗函数 $g(t)$ 唯一决定的, 所以 2δ 是一个定值)。在小波变换中, 变换系数 $W_f(a, b)$ 主要依赖于信号在 $[b-a\Delta\psi, b+a\Delta\psi]$ 中的情况, 时间宽度是 $2a\Delta\psi$, 该时间宽度是随着尺度 a 变化而变化的, 所以小波变换具有时间局部分析能力。
- (6) 若用信号通过滤波器来解释, 小波变换与短时傅立叶变换的不同之处在于:

对短时傅立叶变换来说, 带通滤波器的带宽 Δf 与中心频率 f 无关; 相反, 小波变换的带宽 f 则正比于中心频率 f , 即

$$Q = \frac{\Delta f}{f} = C \quad C \text{ 为常数} \quad (3.9)$$

也即滤波器有一个恒定的相对带宽, 称之为等 Q 结构 (Q 为滤波器的品质因数, 且有 $Q = \text{中心频率}/\text{带宽}$)。

3.4 离散小波变换

在实际应用中, 我们不能用计算机来实现连续小波变换。因此, 需要将其离散化。离散小波变换 DWT(Discrete Wavelet Transform)采用对数函数来量化尺度参数, 而时间参数则依赖于尺度参数, 对于不同的应用, 对数的底可以取不同的数, 最常用的是取 2。如果采用以 2 为底的对数变化, 则在时间轴上的采样速率也以 2 为因子进行变化。在式 (3.5) 中, 假定尺度参数取 $a = 2^{-m}$, 平移参数取 $b = na = n \times 2^{-m}$, 其中, $m, n \in Z$, 则

$$\psi_{m,n}(t) = 2^{m/2} \psi(2^m t - n) \quad (3.10)$$

由此可定义离散小波变换:

$$C_{m,n} = \int_{-\infty}^{\infty} f(t) \psi_{m,n}^*(t) dt \quad (3.11)$$

根据式 (3.10), 可以得到, 当我们调节尺度参数时, 随着 m 变小, 尺度 a 随之变小, 相应的频率变高, 可以观察到更多的细节。若我们调节平移参数, 随着 n 变化, 可以分析不同时段的信息。

3.5 多分辨率分析与 Mallat 算法

多时间尺度分析(多分辨率分析)的思想是指: 小波变换系数 $\omega_f(a,b)$ 能反映函数 $f(t)$ 随时域参数 b 和频域参数 a 变化的特性, 当 a 较小时观察的视野窄, 对时域的分辨率较高, 而对频域的分辨率较低, 可以作细节观察; 当 a 较大时观察的视野宽对时域的分辨率较低, 而对频域的分辨率较高, 可以作概貌的观察。但不同的 a 值下分析的品质因数(中心频率与带宽的比值)却保持不变, 这种多尺度分析的思想也称多分辨率分析思想。

设 V_m 为分辨率为 2^m 时对 $L^2(R)$ 的逼近, A_m 为分辨率为 2^m 时 V_m 上的正交投影算子, 则 $f_{am}(x) = A_m f(x)$ 为在分辨率为 2^m 时对 $f(x)$ 的最佳逼近。令 W_m 为 V_m 在 V_{m+1} 上的正交补, $V_m \perp W_m$, 则:

$$V_{m+1} = V_m \oplus W_m \quad (3.12) \text{ 令}$$

D_m 为分辨率为 2^m 时 W_m 上的正交投影算子, 则 $f_{dm}(x) = D_m f(x)$ 为在分辨率为 2^m 时

$f(x)$ 的细节信号，有：

$$f_{am+1}(x) = f_{am}(x) + f_{dm}(x) \quad (3.13)$$

1987 年，Mallat 巧妙的将多分辨率分析的思想引入到小波分析中，建构了小波函数的构造及信号按小波转换的分解及重构。这种基于滤波器组的小波分解和小波重构算法称为 Mallat 算法^[8]。

$$\text{考虑函数 } \varphi(x), \text{ 令 } \varphi_{m,n}(x) = 2^{m/2} \varphi(2^m x - n) \quad m, n \in Z \quad (3.14)$$

若 $\{\varphi_{m,n}(x), n \in Z\}$ 构成 V_m 的正交归范基，则称 $\varphi(x)$ 为多分辨率逼近的 $\{V_m, m \in Z\}$ 的尺度函数。相应地， $f(x)$ 可分解到尺度函数 $\varphi_{m,n}(x)$ 和小波函数 $\psi_{m,n}(x)$ 上形成正交投影，分别对应分辨率为 2^m 时的逼近信号和细节信号。

一维小波分析的分解和重构公式

由于 $\{\varphi_{m,n}(x), n \in Z\}$ 构成 V_m 的正交归范基， $f(x)$ 在 V_m 上的投影 $A_m f(x)$ 可写为：

$$A_m f(x) = \sum_n a_m \varphi_{m,n}(x) \quad \text{并由此可推出：}$$

$$a_m(n) = \sum_k h(k-2n) a_{m+1}(k) = \sum_k \tilde{h}(2n-k) a_{m+1}(k) \quad \tilde{h}(n) = h(-n) \quad (3.15)$$

$$h(n) = \sqrt{2} \int_R \varphi(x) \varphi(2x-n) dx \quad (3.16)$$

其中 \tilde{h} 为与小波对应的低通分解滤波器， $a_m(n)$ 为 $f(x)$ 的离散逼近信号。

同理， $f(x)$ 在 W_m 上的投影 $D_m f(x)$ 可写为： $D_m f(x) = \sum_n d_m \psi_{m,n}(x)$ ，可推出：

$$d_m(n) = \sum_k g(k-2n) a_{m+1}(k) = \sum_k \tilde{g}(2n-k) a_{m+1}(k) \quad \tilde{g}(n) = g(-n) \quad (3.17)$$

$$g(n) = \sqrt{2} \int_R \psi(x) \varphi(2x-n) dx \quad (3.18)$$

其中 \tilde{g} 为与小波对应的高通分解滤波器， $d_m(n)$ 为 $f(x)$ 的离散逼近信号。

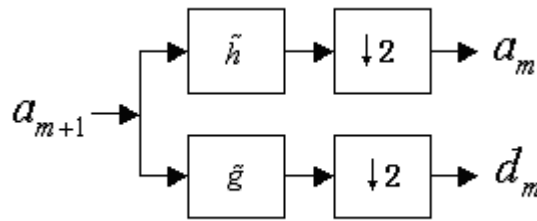


图 3.1 一维信号小波分解

Fig 3.1 1-D Wavelet Transform Decomposition

相应的重构公式为：

$$a_{m+1}(n) = \sum_k h(n-2k) a_m(k) + \sum_k g(n-2k) d_m(k) \quad (3.19)$$

3.6 二维离散小波变换

由于图像是二维信号，因此首先要把小波变换由一维推广到二维。令 $f(x_1, x_2)$ 表示一个二维信号 x_1, x_2 分别是其横坐标和纵坐标， $\psi(x_1, x_2)$ 表示二维的基本小波。对应的尺度函数 $\varphi(x_1, x_2)$ ，若尺度函数是可分离的，即

$$\varphi(x_1, x_2) = \varphi(x_1)\varphi(x_2) \quad (3.20)$$

令 $\psi(x_i)$ 是与 $\varphi(x_i)$ 对应的一维小波函数，则二维的二进小波可表示为以下三个可分离的正交小波基函数：

$$\psi^1(x_1, x_2) = \varphi(x_1)\psi(x_2) \quad (3.21)$$

$$\psi^2(x_1, x_2) = \psi(x_1)\varphi(x_2) \quad (3.22)$$

$$\psi^3(x_1, x_2) = \psi(x_1)\psi(x_2) \quad (3.23)$$

这说明在可分离的情况下，二维多分辨率可分两步进行。先沿 x_1 方向分别用 $\varphi(x_1)$ 和 $\psi(x_1)$ 做分析，把 $f(x_1, x_2)$ 分解成平滑和细节两部分，然后对这两部分再沿 x_2 方向用 $\varphi(x_2)$ 和 $\psi(x_2)$ 做同样分析。

二维小波分析的分解算法为：

$$a_j(n, m) = \sum_l \sum_k h(2m-l)h(2n-k)a_{j+1}(k, l) \quad (3.24)$$

$$d_j^1(n, m) = \sum_l \sum_k g(2m-l)h(2n-k)a_{j+1}(k, l) \quad (3.25)$$

$$d_j^2(n, m) = \sum_l \sum_k h(2m-l)g(2n-k)a_{j+1}(k, l) \quad (3.26)$$

$$d_j^3(n, m) = \sum_l \sum_k g(2m-l)g(2n-k)a_{j+1}(k, l) \quad (3.27)$$

重构算法为：

$$\begin{aligned} a_{j+1}(k, l) = & 4 \sum_k \sum_l h(n-2k)h(m-2l)a_j(k, l) + 4 \sum_k \sum_l g(n-2k)h(m-2l)d_j^1(k, l) \\ & + 4 \sum_k \sum_l h(n-2k)g(m-2l)d_j^2(k, l) + 4 \sum_k \sum_l g(n-2k)g(m-2l)d_j^3(k, l) \end{aligned} \quad (3.28)$$

3.7 提升算法

传统的小波变换是在频率域进行的，其基本的变换工具是傅立叶变换，又称为第一代小波变换。由上节的讨论可知，其变换过程主要是对图像行列分别滤波，进行卷积运算，其过程复杂，运算量大，不利于硬件和实时实现。此外，原始图像的整数数据经过传统小波变换后，得到的是浮点数。从计算机有舍入误差的角度来看，通常这样的小波变换不能无失真地重构，因此不能用于无损图像编码。而且，在对变换后的系数编码压缩时需要先进行量化，量化的结果损失了一定的数据信息，无法完全重构。1994年, Swelden^{[40][41]}提出了一种不依赖于傅立叶变换的新的的小波构造方法—提升方法。提升方法则摆脱了传统的滤波器和傅立叶频域的概念，可以直接利用时域信号，从预测和更新的角度来设计算法。对提升算法

稍加改进,就可以实现整数到整数的变换,从而实现图像的真正无损压缩。所以,基于提升方法的小波变换被称为第二代小波变换。

二维离散小波变换最有效的实现方法之一是采用 Mallat 的塔式分解方法,通过在图像的水平 and 垂直方向交替采用低通和高通滤波得到,这种传统的基于卷积的离散小波变换计算量大,对存储空间的要求高,提升小波的出现有效地解决了这一问题。提升算法相对于 Mallat 算法而言,是一种更为快速有效的小波变换实现方法,它不依赖于傅里叶变换,完全在空域完成了对双正交小波滤波器的构造。Daubechies 已经证明^[42],任何离散小波变换或具有有限长滤波器的两阶滤波变换都可以被分解成为一系列简单的提升步骤,所有能够用 Mallat 算法实现的小波变换,都可以用提升算法来实现。

3.7.1 提升算法的过程

小波变换的基本思想是利用信号间存在的相关性,建立一种信号的稀疏表示。传统的小波变换即第一代小波变换是通过傅立叶变换建立时频分析,小波变换的提升方法则通过欧几里德算法来寻找低通、高通滤波器两者的最大公因子,最大限度地挖掘两者之间的计算冗余,大大降低 Mallat 算法的计算复杂度。其通过基本的多项式插值来获取信号的高频分量,之后通过构建尺度函数来获取信号的低频分量,基本思想就是通过一个母小波(lazy wavelet)逐步构建出一个具有更加良好性质的新的小波,完全在信号的空间域,对信号实施分裂(split)、预测(predict)和更新(update)三大步骤,完成对信号的频率分解。

提升算法给出了双正交小波简单而有效的构造方法,使用了基本的多项式插补来获取信号的高频分量,之后通过构建尺度函数来获取信号的低频分量。提升算法的基本思想在于通过一个基本小波,逐步构建出一个具有良好性质的新的小波,这就是提升的基本含义。

由提升方法构成的小波变换的分裂、预测和更新三个步骤可具体描述如下:

- ① 分裂。将信号 s_j 分裂成为两个互不相交的子集 s_{j-1} 和 d_{j-1} ,常用的算法是将输入信号按奇偶分成两个子集,

$$F(s_j) = (s_{j-1}, d_{j-1}) \quad (3.29)$$

② 预测。针对数据间的相关性,可用 s_{j-1} 去预测 d_{j-1} 。故可采用一个与数据集结构无关的预测算子 P ,使得 $d_{j-1} = P(s_{j-1})$,这样就可以用子数据集 s_{j-1} 代替原始的数据集 s_j 。若用子集 d_{j-1} 与预测值 $p(s_{j-1})$ 的差值去代替 d_{j-1} ,则此差值反映了两者的逼近程度。如果预测是合理的,则差值数据集所包含的信息比原始子集 d_{j-1} 包含的信息要少得多。预测过程的表达式如下

$$d_{j-1} = d_{j-1} - P(s_{j-1}) \quad (3.30)$$

- ③ 更新。由于分解成子集,原来集合的一些特征丢失,通过更新使子集的数

据和原来集合的数据保持相同的特征。也就是通过算子 U 产生一个更好的子数据集 s_{j-1} ，使之保持原数据集 s_j 的一些特性。 s_{j-1} 的定义如下

$$s_{j-1} = s_j + U(d_{j-1}) \quad (3.31)$$

重构数据时的提升公式与分解公式相同，改变计算次序和符号就可以了

$$s_{j-1} = s_j - U(d_{j-1}) \quad (3.32)$$

$$d_{j-1} = d_j + P(s_{j-1}) \quad (3.33)$$

$$s_j = \text{Merge}(s_{j-1}, d_{j-1}) \quad (3.34)$$

其中 Merge 是合并的意思，是将分裂后的子集 s_{j-1} 和 d_{j-1} 重构成初始信号 s_j 。图.2 给出了利用第二代小波变换分解和重构的示意图,利用不同的预测算子 P 和更新算子 U 可以建立不同的小波变换。

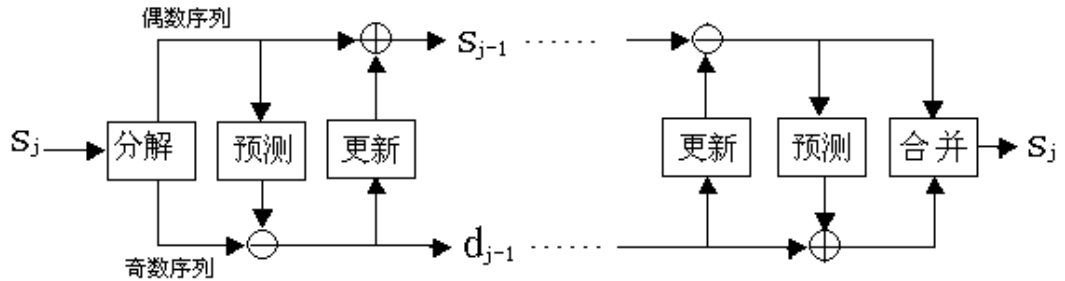


图 3.2 提升算法的框图

Fig 3.2 Block diagram of Lifting Algorithm

3.7.2 提升算法的实现

目前许多人做了将传统小波用提升算法来实现，可以参考文献[6][7][43][44]。经过大量的测试，比较了各种滤波器，浮点型的 Daubechies(9,7)滤波器和整数型的 Daubechies(5,3)滤波器的在图像压缩中综合表现最好。JPEG2000 就选用了这两种滤波器。

Daubechies(5,3)滤波器：（公式中的 $[]$ 表示取整运算）

$$\text{分解算法} \quad s_0[n] = x[2n] \quad d_0 = x[2n+1] \quad (3.35)$$

$$d[n] = d_0 - [1/2 (s_0[n+1] + s_0[n])] \quad (3.36)$$

$$s[n] = s_0 + [1/4 (d[n] + d[n-1])] + 1/2 \quad (3.37)$$

$$\text{重构算法} \quad s_0[n] = s[n] - [(d[n] + d[n-1])/4 + 1/2] \quad (3.38)$$

$$d_0[n] = d[n] + [(s_0[n+1] + s_0[n])/2] \quad (3.39)$$

$$x[2n] = s_0[n] \quad x[2n+1] = d_0[n] \quad (3.40)$$

Daubechies(9,7)滤波器: (公式中的[]表示取整运算)

$$\text{分解算法} \quad s_0[n] = x[2n] \quad d_0 = x[2n+1] \quad (3.41)$$

$$d_1[n] = d_0 - [1/128 (203(-s_0[n+1]-s_0[n])) + 1/2] \quad (3.42)$$

$$s_1[n] = s_0 + [1/4096 (217(-d_1[n]-d_1[n-1])) + 1/2] \quad (3.43)$$

$$d[n] = d_1[n] + [1/128 (113(s_1[n+1]+s_1[n])) + 1/2] \quad (3.44)$$

$$s[n] = s_1[n] + [1/4096 (1817(d_1[n]+d_1[n-1])) + 1/2] \quad (3.45)$$

重构算法略, 由分解算法可推出。

3.7.3 提升算法与传统小波变换的比较

与第一代小波变换相比, 两代小波变换都具有多分辨率的特性, 但是第一代小波变换后的系数是浮点数, 由于受到计算机数据存储和传输方式的限制, 往往不能精确恢复信号, 同时浮点数在压缩编码时, 一般需要进行取整量化, 因此无法达到无损压缩的目的。而第二代小波变换前后的系数都是整数形式, 这不但有利于计算机的存储, 而且可以实现图像的无损压缩。其次, 第一代小波变换后, 数据的重构结果还受到编码时边界数据延拓方式的影响, 而第二代小波变换由于是无损重建信号, 则无须考虑数据的延拓方式; 一代二维离散小波变换最有效的实现方法之一是采用 Mallat 的塔式分解方法, 通过在图像的水平 and 垂直方向交替采用低通和高通滤波得到, 这种传统的基于卷积的离散小波变换计算量大, 对存储空间的要求高, 提升小波的出现有效地解决了这一问题。二代提升算法除具有速度快的优点外, 还能实现任意图像尺寸的小波变换, 一代二进制小波变换要求图像的长宽为 2 的幂数。二代提升算法不依赖傅立叶变换实现小波构造, 可直接在时域中完成小波构造; 它对内存的需求量小, 在当前位置即可进行小波变换, 便于用 DSP 芯片实现; 它不受边界延拓方式的影响等。

在文献[31]中做了二代小波提升算法和 Mallat 的塔式分解方法的比较, 测试图像选用了 Lena 和 Baboon, 这里选用了文献[45]的一个实验结果, 其运算速度比较结果见表 3.1:

表 3.1 小波变换和逆变换的时间

Table 3.1 Process time of wavelet transform and reverse wavelet transform

次数	小波分解 (s)		小波重构(s)	
	Mallat	提升算法	Mallat	提升算法
2	2.906	1.047	2.812	0.984
3	3.234	1.313	3.844	1.219
4	3.781	1.344	3.875	1.266
5	4.234	1.343	3.703	1.250

由实验结果可见,提升算法实现小波变换的时间仅相当于 Mallat 算法的时间的 36%,提升算法的运算速度大大高于传统小波变换,有利于实时处理。

但是 Michael D.Adams 发现传统小波变换更能使能量集中于低频,能量在低频越集中,压缩效果就越好,对比实验结果见文献[45]。但提升算法可实现整数小波变换,免去了量化所带来的失真效应,弥补了其能量不如传统小波变换集中的缺点。越来越多的研究人员在图像压缩中使用提升算法,文献[6][7][46][47]采用提升算法实现小波变换,取得很好的压缩性能。

3.7.4 基于提升方法的小波变换

我们用 Daubechies(5,3)滤波器实现了对 Lena, Barbara 和 Baboon 图像的小波变换。Lena, Barbara 和 Baboon 均为 512×512 的灰度图像,像素值界于 0 到 255 之间。图 3.3 显示了对 Lena 采用提升算法实现两级小波变换后的效果。变换后的图像左上角为低频子带,其余为高频子带,从图中可见,经过变换后的高频子带基本上为黑色,即能量很低,能量被集中到了低频子带。



图 3.3 提升法实现 Lena 小波变换

Fig 3.3 Wavelet transform lena using lifting

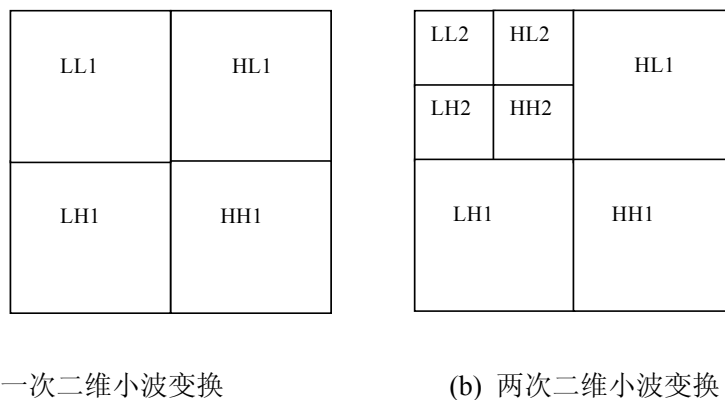


图 3.4 对图像进行二维小波变换

Fig 3.4 2D wavelet transforming image

对经过变换后各子带小波系数的平均值进行了统计(各子带分布见图 3.4)，这些平均值反映了小波变换后能量在各子带的分布，统计结果见表 3.2。

表 3.2 不同子带小波系数的平均值

Table 3.2 Average wavelet coefficient magnitudes in different subbands

	LL2	HL2	LH2	HH2	HL1	LH1	HH1
lena	135.43	5.75	3.92	5.26	3.43	2.43	2.76
barbara	99.17	12.32	6.60	11.96	11.28	3.64	5.02
baboon	130.47	16.06	20.95	22.95	10.89	17.25	13.06

由上面的数据可见，经采用提升法实现小波变换后，低频子带的系数平均值远远高于高频子带，图像的能量主要集中在低频子带，为后面的压缩提供了条件。

4 基于嵌入式思想的小波系数编码

4.1 引言

图像经过小波变换后,并没有实现压缩,只是对整个图像的能量进行了重新分配,小波变换编码就是在小波变换的基础上,利用小波变换的这些特性,对变换后的系数进行合理有效的量化,达到高效压缩的目的。对小波系数进行量化的过程就是对小波图像系数进行组织编码,从而实现压缩的过程。因此研究小波图像系数的有效组织量化方法,是小波变换编码的关键技术之一。

针对小波图像的特点,现有的对小波系数的量化方法主要集中在标量量化和矢量量化两种方法,本章主要研究标量量化方法中最经典的方法--零树编码,或称为嵌入式零树小波编码。零树编码的基本思想是,将小波变换系数的编码分为两部分:一是对用来表明系数分布特性的系数重要性图进行编码;二是对重要系数的幅值进行编码。利用不同尺度的小波系数间存在较强的相关性,将多数的零系数组织成一种树型结构,从而提高了总体编码效率。这种算法得到的比特流中的比特是按重要程度排序的,可方便地实现图像可分级编码与传输。使用这种编码方法,编码者能够在任意一点结束算法,以达到目标比特率或目标失真程度,而这时仍能较准确的重构原来的图像。

4.2 嵌入式编码

嵌入式编码就是指编码器可以生成大小不同的码流文件,而小的文件(有 m 比特),与大的文件的前 m 比特是完全相同的。这样一来,就只需要一个单独的码流就能够以任意低于或等于编码码率的码率来进行解码的编码方式,其主要目的就是要实现渐近的图像传输。我们熟悉的 GIF 图像编码格式就支持渐近传输。这样做的好处是图像的传输可以在任意的时候终止,不需要传输完整的一幅图像,用户就可以看到图像的部分信息。在传输信道很差的情况下,如果图像传输异常中断时,用户也可以根据接收到的数据了解整幅图像的大致信息。而不象非渐近传输那样,如果异常中断传输,用户得到了已传输部分的完整内容,而对未传输部分却一点也没有了解到。

对要研究的图像中的点或像素,用 $p_{i,j}$ 来表示, (i,j) 就是该点的坐标。本文研究的图像都是数字图像,为了简单起见,这里和后面讨论的都是灰度图像,即 $N \times N \times m$ bpp(比特每像素)的图像, N 代表了图像的水平 and 垂直纬度, m 代表表示每个像素所用的比特数。

在渐近传输方案中,一个主要的任务就是选择最重要的信息,它可以使失真

减少最大，把这最重要的信息最先传输。为能做出选择，我们用最小均方误差（MSE, Mean Squared-Error）来衡量失真：

$$D_{MSE}(P - \hat{P}) = \frac{\sum_i \sum_j (P_{i,j} - \hat{P}_{i,j})^2}{N^2} \quad (4.1)$$

其中， P 代表原图像， \hat{P} 为传输后恢复的图像。

从式 (4.1) 可以看出，如果图像数据不能准确的传输到解码端，那么 $|P_{i,j} - \hat{P}_{i,j}|$ 越小，误差越小。如果每个像素的重要性用它的幅度值来表示，那么就可以说，同一像素，其高位的比特比低位比特重要。对于一个 8bit 的像素，如果只允许传送 1bit 的话，显然，传送最高位 bit 比传输最低位比特的效果好。

假设像素按照用二进制表示它的大小所需的最小的比特数来排序。那么，排序就是一个一一对应的映射，设为 $\eta: I \rightarrow I^2$ 。（ η 表示映射， I 表示像素空间， I^2 表示编码空间）

	Sign	S	S	S	S	S	S	S	S	S	S	S	S	S	S
Msb	5	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	4	—	→	1	1	0	0	0	0	0	0	0	0	0	0
	3	—	—	—	→	1	1	1	1	0	0	0	0	0	0
	2	—	—	—	—	—	—	—	→	1	1	1	1	1	1
	1	—	—	—	—	—	—	—	—	—	—	—	—	—	→
Lsb	0	—	—	—	—	—	—	—	—	—	—	—	—	—	→

图 4.1 按幅度大小排列的像素的二进制表示

Figure 4.1 Binary representation of the magnitude-ordered coefficients

图 4.1 是一些按幅度大小排列的像素的二进制表示。最上面一行是数据的符号，最上面的比特是最高位，也即是最重要的。渐近传输就按行进行，从左到右。因为数据已经是递减排列，所以每一列开始的“0” bit 和第一个“1”比特都不需要传送，它们可以自动的在解码端得到恢复。真正需要传送的比特只是图 4.1 中箭头标注的比特。采用嵌入式编码进行渐近的传输，解码端可以以任意低于或等于编码码率的码率来对一个单独的编码文件进行解码，从而为编码提供了更多的灵活性。从上面的分析可以看出，这样的编码方式仿佛是嵌入到了数据比特流内，可以在任意点进行截断，所以就称为嵌入式的比特面编码，或简称嵌入式编码。

4.2.1 嵌入式零树编码

J M Shapiro 提出的嵌入式零树小波编码算法充分利用了图像的小波变换系数各子带之间的相关性,而计算实现却很简单,所以立即引起了许多研究者的巨大兴

趣和进一步探索。零树编码算法的基本思想是采用了零树(Zerotree)这一数据结构来表示小波系数时,充分利用小波变换的优秀时频局域性和挖掘变换系数各子带之间的相关性。该方法有以下几个突出的优点:

- (1) 可对编码后的输出码流进行精确的控制且产生嵌入式码流;
- (2) 由于在每一轮量化扫描时对各子带使用同一个阈值,因此有利于将变换系数大的图像边缘部分保留;
- (3) 容易实现累进传输;
- (4) 不需要做各子带的比特分配和量化步长设计,简化了实现,提高了速度。

经过变换后的图像各子带系数呈现树状的层次分解结构,尺度间有明显的相似性与方向性这些特征,表现为低频系数比较小则其相对应的高频系数一般也比较小。J M Shapiro 提出的嵌入零树小波编码正是因为充分利用了这种关系,获得了较高的编码效率。同时该编码能产生嵌入式码流,支持多码率。

4.2.2 零树编码的概念

有关零树编码的概念如下:

(1) 小波树

一幅图像经过一次二维小波变换后的系数分成四个子带:水平高频垂直高频子带 HH1,水平高频垂直低频子带 HL1,水平低频垂直高频子带 LH1,水平低频垂直低频子带 LL1,如图 3.4(a)所示。对水平低频垂直低频子带 LL1 的系数可以再进行相同的小波变换,如图 3.4(b)所示。

多次小波变换后,其系数形成一个树状结构,称为小波树,如图 4.2 所示。低频子带的系数称为父系数,和它的位置对应的高频子带的系数称为后代系数。树根是最低频子带的结点,它有 3 个子女,其余的子带(最高频子带除外)的结点都有 4 个子女,子结点又有子女,称为孙结点。

小波系数的分布特点是越往低频子带系数值越大,包含的图像信息越多,如图 3.4 (b) 中的 LL2 子带。而越往高频子带系数值越小,包含的图像信息越少。就是在数值相同的情况下,由于低频子带反映的是图像的低频信息,对视觉比较重要,而高频子带反映的是图像的高频信息,对视觉来说不太重要。这样对相同数值的系数选择先传输较低频子带系数的重要比特,后传输较高频子带系数的重要比特。正是由于小波系数具有这些特点,非常适合于图像的嵌入式编码算法,所以在新一代静止图像编码标准 JPEG2000 中也采用小波变换作为图像编码的变换方法。

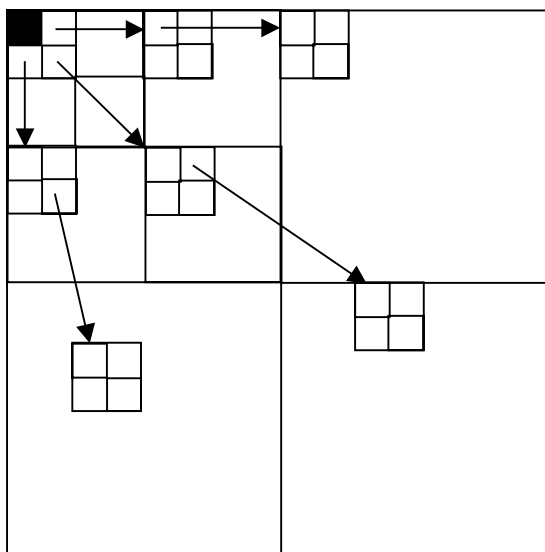


图 4.2 小波树

Fig 4.2 Wavelet tree

(2) 阈值

零树编码是一种嵌入式编码，因为大的系数的丢失所引起的误差比小系数丢失所造成的失真严重，通过先传输比较大的系数，再传小的系数可以减少误差损失，当传输的小的系数不断增加，图像的细节信息也越来越多，从而实现了图像渐进式传输。零树编码中设定一个值，大于等于这个值的系数为重要系数，优先传输，设定的这个值称为阈值 K （门限）。为便于操作，门限一般设为 2 的多次方。

(3) 不重要系数

如果小波系数的幅度小于阈值 K ，即 $|x| < K$ ，则小波系数相对于阈值 K 是不重要（次要）的。对于不重要的系数视为“零”，这实质上是一种标量量化。由于零及零附近系数的概率处于曲线的峰值位置，经过阈值处理后，零（以下将不重要的系数称为零）的个数非常大，零树正是一种把这些零如何组织起来的一种新的数据结构，零树方法是对零如何进行编码的一种方法。

(4) 零树

零树是基于小波系数的相关性的一种假设：如果在低分辨率高频子带上的小波系数相对于阈值 K 是不重要的，那么位于同方向同空间位置高分辨率子带上的所有小波系数相对于阈值 K 也是不重要的。把满足这种假设的系数用树状结构表示出来，就是零树。一幅图像的小波变换系数中，有多棵大小不一的零树。

通过对实际图像的小波分解发现，当一个父系数是不重要的，那么其所有的后代系数也很可能是不重要的。这实际反映了系数各子带之间的相关性。用一个符号就能表示许多系数（父系数和它的后代系数），从而可以实现高效率压缩。

(5) 零树根

零树中位于最低分辨率的那一个零节点称为零树根，它满足：1、该点为零；2、该点不是最高分辨率子带的节点；3、所有各高分辨率子带的节点(即它的所有后代)都为零。所以对这棵零树，我们只需对零树根进行编码，用 T 表示零树根，其实质表示了一个许多零节点组成的集合。

(6) 重要系数

重要系数表示一个系数相对于阈值有意义的，即 $|x| \geq K$ 。若系数为正，则用 P 表示；为负，则用 N 表示。

(7) 孤立零点

孤立零点表示一个系数相对于阈值是不重要的，但其后代中有重要系数。用 Z 表示。

4.2.3 零树编码算法原理

(1) 扫描顺序

EZW 在编码的时候输出的位置信息隐含在特定的扫描顺序里，解码器使用和编码器相同的规律扫描解码，就可以恢复出原始的图像。扫描的顺序如图 4.3 所示。这样的扫描顺序能保证粗尺度系数在细尺度系数前被扫描，当前子带中的所有系数扫描完后才进入下个子带，也即低频信号优先编码，从而实现了嵌入式编码。

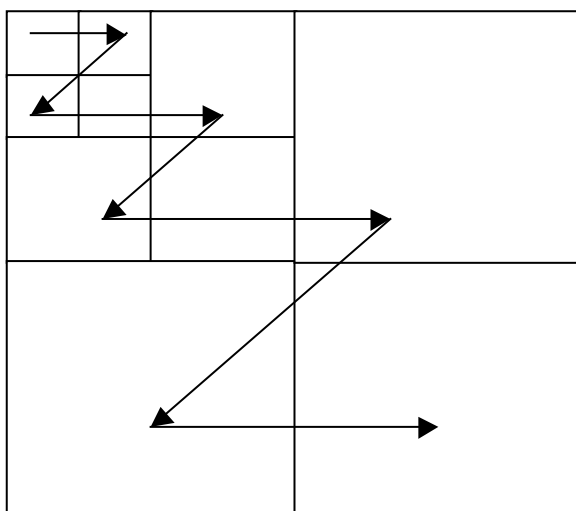


图 4.3 EZW 的扫描顺序

Fig 4.3 Scan order of EZW

(2) EZW 编码

EZW 的编码就是按照扫描的顺序，相应于给定的阈值 K ，将每个小波系数分

成 POS(简写为 P)、NEG(简写为 N)、IZ(简写为 Z)和 ZTR(简写为 T)的码流。EZW 编码算法在一阈值下经过两个过程：主过程 (Dominant_pass) 和辅过程 (Subordinate_pass)。然后将阈值降为原来的一半，步骤同上。当阈值不断降低，就实现了逐次逼近的嵌入式编码。其过程可以描述为：

```

K = K0;
Do{
    Dominant_pass(image);
    Subordinate_pass(image);
    K = K/2;
}while(K>0);

```

其中主过程实现以下 3 项任务：

- ① 按一定的顺序扫描小波系数，根据小波系数输出码流，设取出的系数为 $x[n]$ ，当前阈值为 K ，其输出规则如下：
 - 如果 $x[n]>K$ ，则输出 P；
 - 如果 $x[n]<-K$ ，则输出 N；
 - 如果 $|x[n]|<K$ ，且该系数后代中有重要系数，则输出 Z；
 - 如果 $|x[n]|<K$ ，且该系数后代中没有重要系数，则输出 T；
- ② 将主表中的重要系数抽取出来，其绝对值放入被称为附表的一维数组中。
- ③ 在主表中重要系数的位置填零。

辅过程对重要系数再进行精细编码，根据附表中存的系数输出 0 或者 1，0 表示系数在下半区，即 K 到 $3/2 K$ 之间，1 表示系数在上半区，即 $3/2 K$ 到 $2K$ 之间。最后当阈值 K 降为 1 时，根据主通道和从通道输出的码流就能无误的表示出所有系数的值。要实现无损压缩，还要对输出的码流进行熵编码，一般采用算术编码或者游程编码。因为各轮的阈值不一样，最后在得到的附表中，列在前面的非零值的点比后面的重要。在判别中只有四种符号：对应于某门限的正值、负值、零树根、孤立零，从而大大压缩了所需的比特数。对应于四种符号再加上附加的非零值的数值范围的符号可以用熵编码按概率的大小进一步予以压缩，达到低比特率的要求。零树编码的另一个特点是各层子带中有意义的系数是按父系到子系、重要的值优先的原则在附表中排列，可以按所需的压缩比截取有意义的值，因此在压缩编码中比较灵活。

解码时根据 P、N、Z、T 码流和扫描顺序就能确定系数的位置，同时由码流和附表的价值就能确定系数的值。

(3) 细化。

细化过程就是逐次量化逼近的过程。用一个阈值序列 K_0, K_1, \dots, K_{N-1} 来判定重要系数。其中,

$$\begin{cases} K_0 = 2^N & (N = [\log_2(\max |x_{i,j}|)]) \\ K_i = K_{i-1} / 2 \end{cases} \quad ([] \text{表示取整运算}) \quad (4.2)$$

随着阈值的逐步降低, 就能越来越精细地表示原图像。当阈值降到 1 时, 就能实现无损压缩。

零树描法编码的流程图如图 4.4 所示。

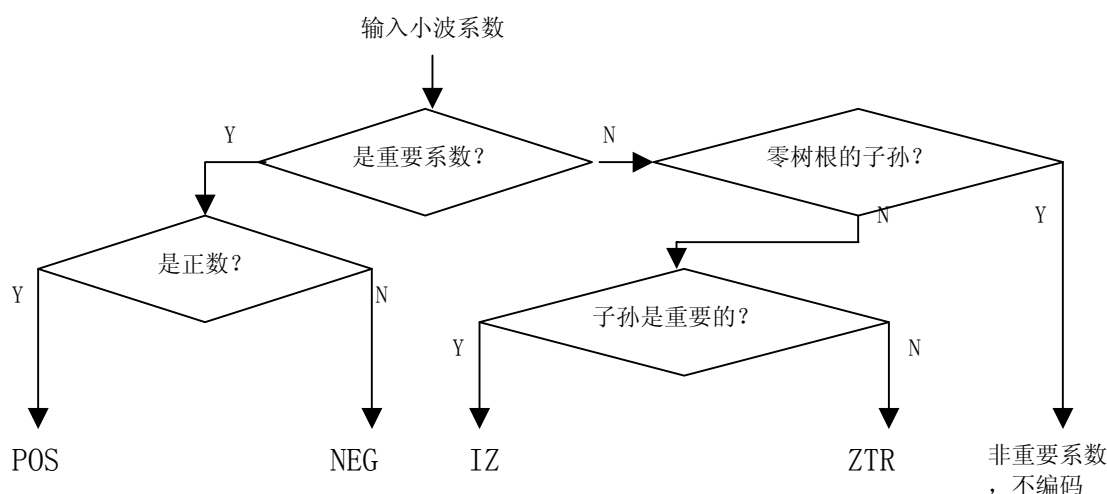


图 4.4 零树描法编码流程

Fig 4.4 Zerotree scanning method coding processing

4.3 分层树集合分割算法 (SPIHT)

Shapiro 提出的 EZW 算法, 对小波系数编码的研究作出了重大的突破, 在基于小波的编码研究中起了里程碑式的作用。但 EZW 算法的编码效率比较低, 经过 EZW 编码后, 数据还需要经过算术编码才会有较好的压缩效果, 否则压缩数据的效果并不理想。文献[46]就是对 EZW 进行改进的算法, 也对 EZW 做了详细阐述。

1996 年, A.Said 和 W.A.Pearlman 根据 Shapiro 的 EZW 算法的基本思想, 提出了一种新的且性能更优的实现方法——分层树集合分割算法 (SPIHT: Set Partitioning In Hierarchical Trees) 算法, 它比此前的 EZW 算法提供了更好的性能, 同样实现了完全嵌入式的编码。

下面就介绍 SPIHT 中的一些有价值的思想和基础概念, 对于 SPIHT 的详细叙述可以参考文献[18]。

4.3.1 集合分类与分级量化

SPIHT 的主要特点之一就是排序数据不是明确的传送的,而是基于如下的事实,即任何算法的执行路径是由在分支点比较的结果决定的。因此,如果编码器和解码器有同样的排序算法,那么如果解码器收到了数据大小的比较结果,它就能复制编码器的执行路径,并且排序信息就能从执行路径中获得。

在这样的排序算法中,不需要对所有的系数进行分类,只需要判断当前系数是重要还是不重要。所以,可以简单的选出满足 $2^n \leq |c_{i,j}| < 2^{n+1}$ 的系数,在每个通道都有给定的 n 值。当给定一个 n ,如果 $|c_{i,j}| \geq 2^n$,那么该系数就是重要的;否则,它就是不重要的。

此时,只需要一个分类算法,把点的集合分成分块的子集 T_m ,并完成数值大小的测试:

$$\max_{(i,j) \in T_m} \{|c_{i,j}|\} \geq 2^n? \quad (4.3)$$

对式(4.3),如果得到否定的回答(即该子集是不重要的),它就可以知道在 T_m 中的所有系数都是不重要的。如果是得到肯定的回答(即该子集是重要的),那么就根据确定的规则再分割 T_m 得到新的子集 $T_{m,l}$,然后,重要性测试又用于这个新的子集。集合划分的过程一直持续下去,直到所有单个坐标重要子集都做了大小测试,以鉴别每个重要的系数。这里,明显的好处就是编码器和解码器可以共享一个分割集合的规则。唯一不同的是编码器用式(4.3)做出判断;而解码器是直接收到式(4.3)的结果,然后用它来做出判断。

可以选择下面的函数来作为判决函数:

$$S_n(T) = \begin{cases} 1, \max_{(i,j) \in T} \{|c_{i,j}|\} \geq 2^n \\ 0, otherwise \end{cases} \quad (4.4)$$

其中, $S_n(T)$ 就表示了坐标为 T 的集合的重要性。这里为简化单个点集合的符号,把 $S_n(\{(i,j)\})$ 写作 $S_n(i,j)$ 。通过 n 来计算阈值的大小, n 将在每一层逐次的变化,从而使阈值不断变化,从而实现了分级量化的重要性测试。从这里也可以看出,这类编码方法将如图 1.1 中量化和编码过程结合在了一起进行。

同时,分级量化的思想也为实现嵌入式的编码打下了基础。参考图 4.1,可以更好的进行理解。

4.3.2 空间方向树

一般情况下,一幅图像的能量大部分都集中在它的低频成份,主要是描述了图像的轮廓;而能量少得多的高频部分主要是描述了图像的一些细节。小波变换

将图像分解成具有不同分辨率和不同方向特性的子带，并使图像信息能量适当集中在变换后图像左上方的某些子带上。

小波变换后的子带构成了一个金字塔形的结构，当从子带金字塔的最高层向最低层运动时，子带间的差异就在减小；而且，在子带间已经观察到了存在一个空间的自相似性，并且如果在子带金字塔中沿同样的空间方向向下运动，就可以期望系数得到更好的大小排序。例如，期望大量不活跃区域被分在子带金字塔的最高层，但它们又被复制到了同一空间方向的较低的层次上。这样，就可以用一个树状结构来表示这种空间关系，这种树状结构就称为空间方向树^[16]，如上一节中图 4.2 所示。

图 4.2 显示了空间方向树是如何定义在一个带有循环的四子带分裂的金字塔结构上的。树上的每个节点对应图像中的一个点，并由点的坐标来识别。金字塔的最高层的称为树的根，也是一个 2×2 的点阵，它的直接子节点分裂规则特殊，因为其四个点中有一个（即黑色的那一个）将没有后代。而其它节点的直接后代（也称子代），相应于树状结构中下一个层次的同一空间方向的点，如图中箭头标示的方向。那么，按照这样的定义方式，每个点要么有四个子代（构成 2×2 的点阵）；要么没有后代，如根处的点或最后层的点。

根据空间方向树的特性，定义下面的集合^[16]：

$O(i, j)$ ：节点 (i, j) 的所有子代的坐标的集合；

$D(i, j)$ ：节点 (i, j) 的所有后代的坐标的集合，即包括子代、孙代、曾孙代……；

H ：所有空间方向的树的根节点（在金字塔的最高层的节点）的坐标的集合；

$L(i, j)$ ： $L(i, j) = D(i, j) - O(i, j)$ ，即除了子代以外的所有后代的集合。

现在，就可以利用用空间方向树来制定划分子集的集合分割规则，分割规则如下：

- (1) 用集合 $\{(i, j)\}$ 和 $D(i, j)$ 来形成初始的划分，所有的 $(i, j) \in H$ ；
- (2) 如果 $D(i, j)$ 是重要的，那么就把它分割成 $L(i, j)$ 和四个单元集合，单元集合的 $(k, l) \in O(i, j)$ ；
- (3) 如果 $L(i, j)$ 是重要的，那么它就被分割成四个 $D(k, l)$ 的集合，其中 $(k, l) \in O(i, j)$ 。

4.3.3 编码方法

SPIHT 的详细算法参见文献 [18] 的算法 II，这里不准备详细的列出原算法，主要讨论其编码方法中的独特之处。

因为 SPIHT 是按次序测试子集的重要性，次序就非常重要，在实际执行时，重要信息存储在三个排序的表中：不重要集合列表（LIS, List of Insignificant Sets）、不重要点列表（LIP, List of Insignificant Pixels）和重要点列表（LSP, List of

Significant Pixels)。在所有的表中，每个记录由坐标 (i, j) 来识别，在 LIP 和 LSP 中坐标代表单独的点，而在 LIS 中代表集合 $D(i, j)$ 或 $L(i, j)$ 。为在上面两个集合之间进行区别，如果一个 LIS 记录代表 $D(i, j)$ 就用字母 D 表示，如果代表 $L(i, j)$ 就用字母 L 表示。

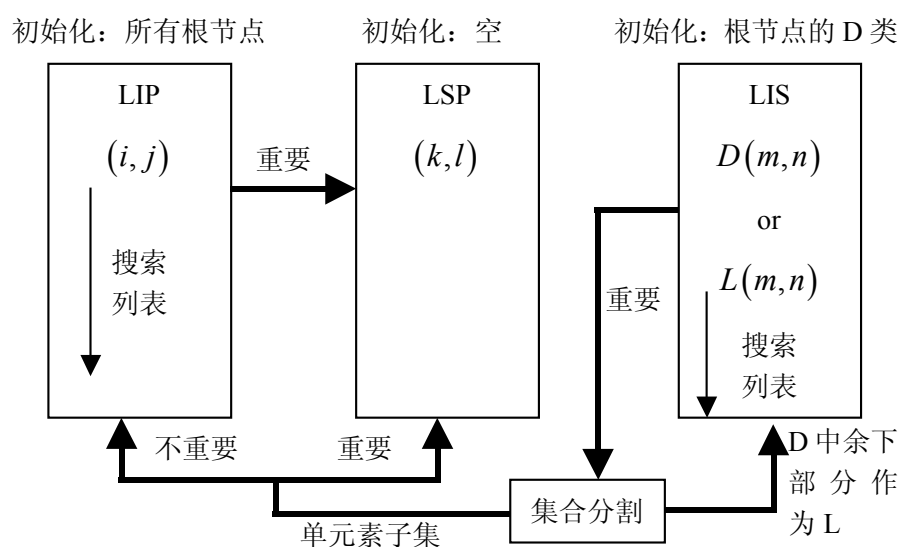


图 4.5 SPIHT 的分类过程

Fig 4.5 the classify process of SPIHT

图 4.5 给出了 SPIHT 的分类过程的流程示意图。前面已经讨论了 SPIHT 的分类算法思想，这里结合 LIP、LSP 和 LIS 三个表来说明其分类过程的执行情况。从图中可以看出，集合的分裂都是在处理 LIS 表时完成的，而在 LSP 表中的点已经完成了排序的过程。那么在 LSP 表中的点实际就是按图 4.1 那样存储的，对 LSP 表的处理就是完成嵌入式编码的过程。

SPIHT 算法的解码器完全复制了编码器在对重要系数分类时的执行路径，这样，为得到想要的解码器算法，就仅仅只需要把算法中的“output”替换为“input”就可以了。当重要系数的坐标被加到 LSP 后面的时候，就可以获得次序信息 $\eta(k)$ 。任何时候解码器输入的数据，它的三个控制列表——LIS，LIP 和 LSP，都和编码器在输出数据的时候用到的是一样的，这就意味着解码器确实从执行的路径中获得了次序的信息。容易看出使用这一方案，编码和解码有同样的计算复杂度。

这样一来，SPIHT 使得集合的表示更为精简，提高了编码的效率。小波系数

经 SPIHT 编码后, 即使不再使用算术编码, 也有很好的压缩效果^[48]。

4.4 嵌入式块集合分裂编码器 (SPECK)

嵌入式块集合分裂编码器 (SPECK: the Set Partition Embedded block coder) 是 Islam 和 Pearlman 在 1999 年发表在[1111]中的小波系数编码算法, Pearlman 也是 SPIHT 的作者之一。同前面提到的 EZW 和 SPIHT 一样, SPECK 算法也是属于分级量化的重要性测试方法这一类的编码方法。但同 EZW、SPIHT 等方法的一个重要区别在于, SPECK 没有使用树状结构来处理小波变换后的子带; 而是利用了块状结构的集合。SPECK 继承了前面的编码方法的特性, 如: 完全嵌入式的编码、实现了渐近的图像传输、低计算复杂度、编/解码器具有相同的结构、适于无损和有损图像压缩等; 而且还具有对动态存储空间要求低的特性, 这是采用块状结构带来的好处。详细的 SPECK 算法参见文献[19], 下面只讨论 SPECK 算法的核心部分。

4.4.1 块状集合分割

在 SPECK 中, 将图像经小波变换后的每个子带都视为一个块, 其中最小的块为 2×2 的块。其仍然采用式 (4.4) 的判决函数, 对每个集合进行判断, 当集合重要时, 就把集合分割成新的集合。其新集合划分的方法仍然是由一个集合分割成四个新的集合, 如图 4.7 所示, 新的集合只有原集合的四分之一大小。

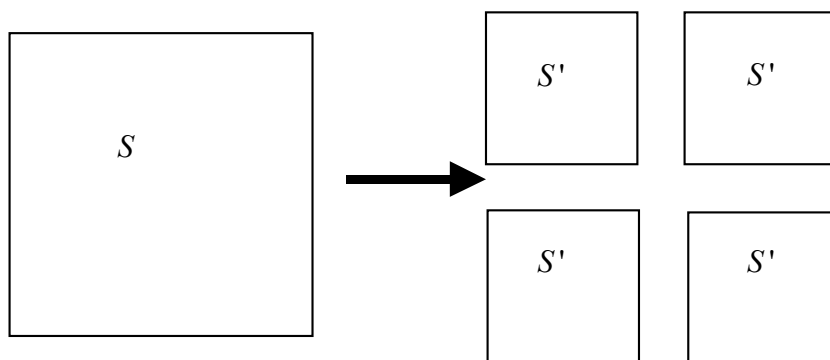
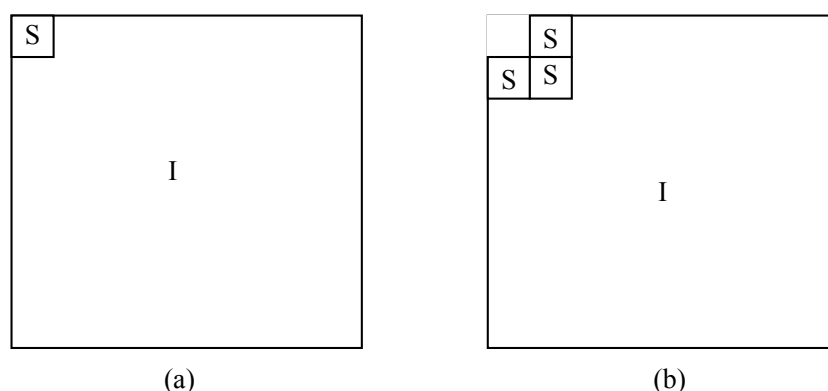


图 4.6 集合分裂方式

Figure 4.6 Partitioning of set

由于没有了树状结构, 小波变换后的系数按块状结构进行后期处理, 如图 4.7 所示。



(a) 原图像划分为 S 和 I 集合；(b) I 集合分裂后的情况

图 4.7 块状结构示意图

Figure 4.7 Diagram of block

从图 4.7 中可以看到，SPECK 简单的把系数划分为了 S 集合和 I 集合。S 集合的尺度可以任意设定，而且在编码过程中也是变化的，一般和变换后的小波子带尺寸相同。

比较前面讨论的 SPIHT 算法，采用这样的块划分的好处是，在 I 集合不重要时只需要 1 个 bit，而 SPIHT 却需要 3 个比特。没有了树状结构，处理起来也很简单。在 I 集合重要时，原 I 集合分裂为 3 个 S 集合和 1 个新的 I 集合，如图 4.7 (b) 所示，这 3 个 S 集合就是当前子带中最低频的块。这主要是考虑到小波系数具有的特点：经小波变化后，能量主要集中在子带的低频部分，即在子带的左上方。那么，满足新的阈值的点更有可能在新产生的 3 个 S 集合中，而不是在新的 I 集合中。这样的划分就可以使含有重要系数可能性大的集合（S 集合）的尺寸比含有重要系数可能性小的集合（I 集合）的尺寸小，而且也会先得到处理。

因为采用了块状结构，在给定的时间内，仅处理一个相关的块，处理完毕后才考虑下一个，这样的方法就节省了动态存储空间。另一方面，SPECK 提高了容错性。因为块间是独立编码的，在传输发生误码时，只有误码所在的块受到影响。而不论是 EZW 还是 SPIHT，由于采用了树状结构，出错时会对整个图像产生影响。

4.4.2 编码方法

与前面的 SPIHT 不同，SPECK 只建立了两个表：LIS 和 LSP。在 SPECK 中，S 集合分裂到最后就只有一个点，它可以视为特殊的 S 集合。这样就可以把 SPIHT 中的 LIP 和 LIS 合并为一个新的 LIS。

SPECK 的编码由四个过程组成，各过程的功能如下：

(1) 初始化

将小波变化后的系数 X 分为两个集合: S 和 $I=X-S$;

计算阈值 $n = \left\lceil \log_2 \left(\max_{(i,j) \in X} |c_{i,j}| \right) \right\rceil$;

LSP 置空, 把 S 加入到 LIS 中。

(2) 分类过程

处理 S 集合: 对于非单点 S 集合, 当 S 集合重要时, 分裂为四个新的 S 集合。搜索新的 S 集合, 判断新的 S 集合的重要性, 不重要就放到 LIS 的末端; 如果重要, 又重复集合的分裂过程。

I 集合: 测试 I 集合重要性, 如果重要, 就分裂为 3 个新的 S 集合和一个 I 集合, 加到 LIS 末端。

(3) 精练过程

处理 LSP, 输出第 n 个重要位。

(4) 量化级修正

即改变阈值 (如: $n \leftarrow n-1$), 然后返回 (2)。

4.5 现有算法的不足

本章前面几节深入讨论了基于嵌入式思想的小波系数编码方法, 阐述了嵌入式编码实现渐近图像传输的原理。着重分析了在这类方法中具有代表性的三种算法: EZW、SPIHT 和 SPECK 算法, 其中 EZW 和 SPIHT 是基于树状结构的, 而 SPECK 是基于块状结构。它们各自都考虑了小波系数的特性, 实现了完全嵌入式的编码、图像的渐近传输等功能, 能够同时支持无损和有损的图像压缩编码。

尽管这些算法有较好的性能, 但都还存在一些不足, 有待进一步的改进和提高。比如 EZW 算法存在以下问题:

(1) 由于编码时形成多棵零树, 因而要多次扫描图像, 造成效率很低, 而且每一棵零树必须在前一棵零树形成之后才能形成, 所以很难用并行算法优化。

(2) 对所有的频域进行等同重要度的编码, 不能充分利用小波变换的特点。

(3) 在一棵零树中包含的元素越多, 则越有利于数据压缩, 在 EZW 算法中存在这样的树间冗余。

(4) 通过对小波系数的分析发现, 在同一子带中相邻元素间有一定的相关性,

SPIHT 采用的树状结构相对 SPECK 的块状结构来说要复杂些, 这就使额外的开销增加, 会影响压缩性能。虽然 SPECK 采用的块状结构简单, 但其采用的分类和搜索策略还有不足, 造成其时间开销过大

本文针对 EZW 算法运算复杂、扫描效率低的缺点做出了相应的改进, 希望能

够进一步提高图像的无损压缩性能，减少算法的时间开销，同时仍保持对有损压缩的支持。在接下来的一章，介绍了本文提出的基于整数小波变换的图像压缩编码方案，其中 5.2 节详细阐述了本文提出的对 EZW 的改进算法——不完全扫描法。

5 基于整数小波变换的图像压缩编码方案

基于前面的讨论和分析, 本文对嵌入式零树小波编码方法进行了一些改进, 提出改进的零树编码方法---不完全扫描法, 并结合基于提升方法的整数小波变换, 设计了一个可分级图像编码系统。该系统包括图像整数小波变换、改进的零树量化和熵编码等几部分, 系统流程图如图 5.1 所示。

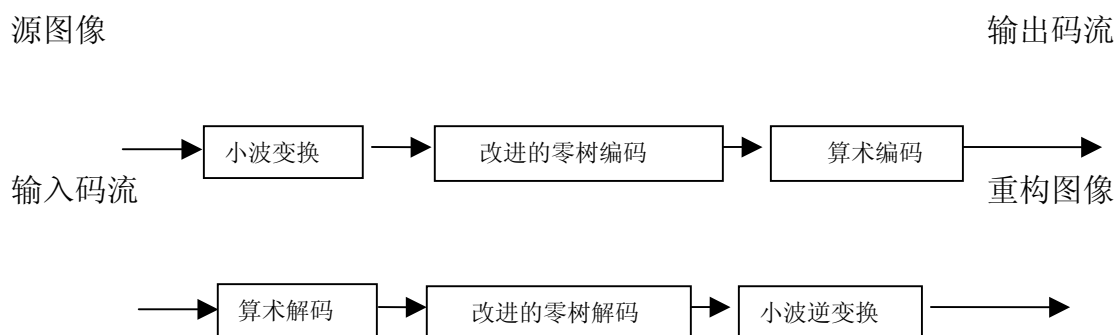


图 5.1 图像编/解码框图

Fig 5.1 graph of image encoding/decoding based on wavelet transform

图像首先进行小波变换, 若采用浮点小波变换还要经过系数量化, 本文采用基于提升算法的整数小波变换, 变换后的系数仍然是整数, 故不需取整。经过变换后的图像能量集中到低频子带, 各子带系数呈现树状的层次分解结构, 尺度间有明显的相似性与方向性这些特征。本文提出的改进的嵌入式零树小波编码---不完全扫描法, 充分利用了这种关系。最后对变换输出的小波系数进行熵编码, 本文选用了 2 阶自适应算术编码。

解码过程为首先对压缩文件进行算术解码, 然后零树解码, 最后经过小波逆变换就能重构原图像。

下面将详细介绍该系统主要部分的原理及其实现过程, 最后给出实验结果。

5.1 基于提升算法的整数小波变换

图像压缩编码时, 小波基的选取是一个关键的问题。目前用于图像处理的小波有多种, 但至今还没有找到一种小波基对于所有类型的图像都适用。同一幅图像, 用不同的小波基进行分解所得到的数据压缩效果是不同的。在应用于图像压缩时, 希望经小波变换后得到的图像在三个方向的细节分量具有高度的局部相关性, 而整体相关性被大部分, 甚至被完全解除。在选取小波基时, 应综合考虑以下三个因素:

- (1) 小波基的正则性与图像数据压缩效果的关系;
- (2) 待处理图像与小波基的相似性;
- (3) 综合考虑压缩效率和计算复杂度, 在两者间做出均衡。

一般而言, 小波函数正则性越高, 图像压缩后重构质量越好; 小波基与待处理图像越相似, 压缩比越高。而且用提升方法构造小波时, 小波消失矩越大, 计算复杂度越高。

本文主要选用了两种小波: 5/3 小波和 9/7-M 小波^{[12][42][45][54]}。5/3 能实现整数小波变换, 结构简单, 变换速度快, 消耗系统资源小, 有利于实时处理和 DSP 应用。9/7-M 很适合图像的无损压缩^{[49][50][51][52][53]}。所以本文选用的两种小波是很有代表性的。各个小波的尺度函数和小波函数见图 5.2~5.3。

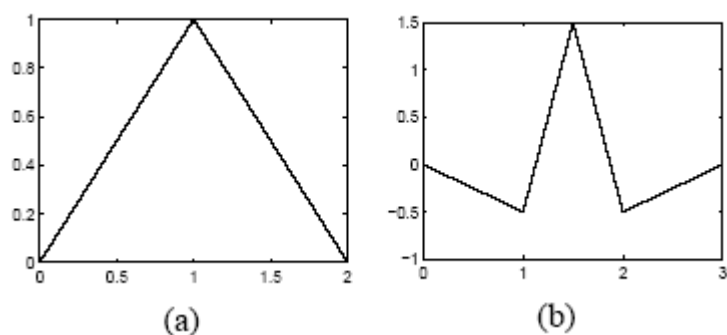


图 5.2 5/3 小波尺度函数 (a) 和小波函数 (b)

Fig 5.2 5/3 wavelet scaling function (a) and wavelet function (b)

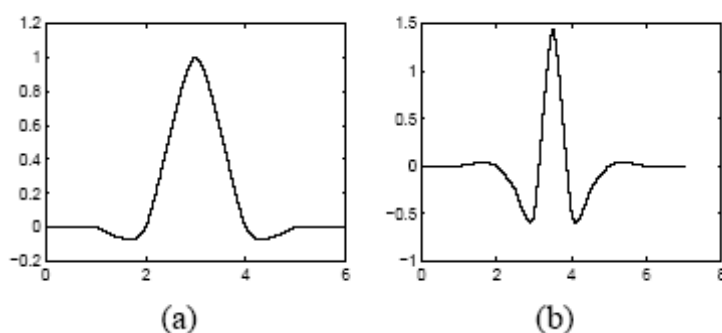


图 5.3 9/7-M 小波尺度函数 (a) 和小波函数 (b)

Fig 5.3 9/7-M wavelet scaling function (a) and wavelet function (b)

下面给出这两个小波按照提升步骤得到的正向变换公式。

5/3 小波:

$$d[n] = d_0[n] - \left\lfloor \frac{1}{2} (s_0[n+1] + s_0[n]) + \frac{1}{2} \right\rfloor \quad (5.1)$$

$$s[n] = s_0[n] + \left\lfloor \frac{1}{4}(d[n] + d[n-1]) + \frac{1}{2} \right\rfloor \quad (5.2)$$

9/7-M 小波:

$$d[n] = d_0[n] + \left\lfloor \frac{1}{16}((s_0[n+2] + s_0[n-1] + 2s[n-1]) - 9(s_0[n+1] + s_0[n])) + \frac{1}{2} \right\rfloor \quad (5.3)$$

$$s[n] = s_0[n] + \left\lfloor \frac{1}{4}(d[n] + d[n-1]) + \frac{1}{2} \right\rfloor \quad (5.4)$$

本文用提升算法代替传统小波变换, 基于如下考虑: 1、提升算法小波变换前后的系数都是整数形式, 这不但有利于计算机的存储, 而且可以实现图像的无损压缩; 2、它不受边界延拓方式的影响; 而一代二维离散小波变换最有效的实现方法之一是采用 Mallat 的塔式分解方法, 通过在图像的水平 and 垂直方向交替采用低通和高通滤波得到, 这种传统的基于卷积的离散小波变换计算量大, 对存储空间的要求高, 提升小波只用几步简单的加法和乘法就可以得到, 运算速度大大提高。

5.2 对嵌入式零树编码算法的改进-----不完全扫描法

小波变换适用于图像压缩的特点已经在前面作了阐述, 但需要再次强调的是小波变换本身并不能压缩图像数据它只是提供了减少比特率的一种可能, 必须结合其它编码技术对小波系数进行编码, 才能实现压缩数据的目的。

在前面部分已经介绍了对小波变换后的系数进行编码的主要方法, 以及它们的发展和研究现状。目前, 对小波系数的编码研究主要利用了小波系数的统计特性, 对小波系数进行合理的组织, 对重要系数采用合理的量化方式, 从而实现压缩数据的目的。本节着重对 J M Shapiro 提出的嵌入式零树小波编码算法存在的不足进行了改进。

EZW 算法被认为是最好的小波系数编码方法, 图像经过小波变换后的系数在不同子带间存在一定的相关性, 即低频子带的系数不重要, 则相对应的高频子带系数为不重要的可能性很大。EZW 用零树来表示这种结构, 用一个符号表示了一个集合, 从而节省了用于表示系数的比特数。这正是 EZW 编码能实现高效压缩的原因通过研究发现, EZW 算法也存在不足, 目前的改进主要集中在提高编码效率上, 采用一定的编码方法提高压缩比, 如文献[47]采用阶梯型量化的方法, 文献[55]采用编码中增加零树的方法。但对降低系统开销和提高编码速度方面的研究则不多。在实时处理和实际应用中, 受环境的限制和处理时间的要求, 往往希望内存占用少, 处理时间快。因此, 寻找一种运算量少、开销低的算法是本文的主要目标。

在一阈值下, EZW 算法有主辅两个编码过程。其中辅过程对重要系数再进行级间排序, 根据附表中存的系数输出 0 或者 1, 0 表示系数在下半区, 即 K 到 $3/2 K$

之间, 1 表示系数在上半区, 即 $3/2 K$ 到 $2K$ 之间。这样增加了操作的复杂度, 同时, 重要的系数必须全部进入附表, 在整个编码过程中要一直保存在内存中。需要占用较多的内存来保存这些重要系数。

级间排序指的是副表中的重要系数被细化编码输出后, 在阈值减半进行下一次扫描前, 将副表中细化后的重要系数按它们所在区间进行排序。在本级阈值扫描编码完成后, 阈值将取为当前阈值的一半, 并对主表中所有不重要系数进行扫描判决, 即主扫描。主扫描完成后, 对当前扫描得到的重要系数及以前扫描得到的重要系数按照先后顺序进行细化编码输出。这样, 本次扫描得到的重要系数、上一级阈值细化比特为 1 的系数经排序后, 将排在上一级细化比特为 0 的系数的前面, 在本级的编码中这些被排在前面的系数的细化比特将被优先编码输出, 幅度大的系数优先编码。

这样每一次副表扫描后, 都要根据当前的阈值对所有重要系数进行排序, 保证把经过细化后的系数进一步排序。然而仔细分析排序的作用, 就会发现它大大增加了运算量, 但带来的压缩效率却是有限的。分析前一章的零树预测过程可以知道, 排序其实只对同一级的系数(即在同一阈值被检测为重要系数)起作用;对不同级间的系数是没有影响的, 因为不同级间的系数的大小是确定的, 先后顺序也是确定的。排序其实只是同级系数由于细化精度不够(编码输出的比特还不越多, 压缩比还较高), 在一定误差范围内进行的。这主要是为了编解码的同步, 因为解码端只能根据主表的节点符号和副表的细化比特来重构系数, 并确定重构系数所在的空间位置。因而只能是细化一级, 误差范围缩小一级, 系数排序一次, 不断进行, 直到编码或解码停止为止, 非常复杂。复杂的排序带来的性能改善非常小。因为排序只是在同级系数之间进行, 只是把同级系数大的排在同级系数小的前面, 并且还不能彻底按照系数本身的大小一次到位, 而只能随着不断的细化逐渐进行。这样排序带来的性能改善主要看同级系数的差别大小来确定。级间系数差别大, 排序带来的改善就大, 反之就小。由此, 排序带来的性能改善主要来自于大阈值得到的重要系数, 小阈值得到的重要系数排序带来的改善则较小。然而, 实际图像的小波系数中总是大幅值的系数非常少, 而小幅值的系数非常多。这样, 当重构图像的质量达到一定水平时, 传输的系数阈值已经非常小了, 系数的误差范围也非常小。例如, 对于 $256 \times 256 \times 8\text{bit}$ 的 Lena 图像, 压缩比为 32 时, PSNR 已小于 28db, 重构图像质量已经较差, 而此时系数的阈值已经为 8, 而误差范围为 $8/2=4$ 。就是说, 此时排序已经只对小于 8 而大于等于 4 的系数有效了, 因为大于 8 的系数肯定要小于 8 的系数传输。如果压缩比更大些, 排序的作用更加不明显。所以, 可以取消级间排序, 以得到算法更快的执行速度。

基于以上分析本文提出了一种改进的嵌入零树编码算法---不完全扫描法。该

算法保持了零树结构,从而保持了 EZW 高效率编码的性能;省略级间排序,在保持嵌入式码流的同时降低了运算操作复杂度,可以大大提高编码速度。不完全扫描法算法如下:

```

K = K0;
Do{
    Dominant_pass(image);
    K = K/2;
}while(K>0);

```

其中主过程 (Dominant_pass) 为:

```

do{
    按一定的扫描顺序取出一个系数 x [n];
    如果 x [n]>K,则输出 POS, x [n] = x [n] - 3/2 T;
    如果 x [n]<-K, 则输出 NEG, x [n] = x [n] + 3/2 T;
    如果|x [n]|<K, 且该系数后代中有重要系数, 则输出 IZ;
    如果|x [n]|<K, 且该系数后代中没有重要系数, 则输出 ZTR;
}while(还有系数没有被扫描);

```

其中扫描的顺序与原算法相同。在任意时刻都可以停止输出码流,则为有损压缩,此时能恢复出图像的大致信号,随着码流的增加,细节信号逐步增多。由于总是优先传送重要系数和低频带的信号,能使图像最大限度地减少失真。当阈值为 1 时,则实现无损压缩。

解码过程为:系数先初始化为零,然后把熵编码的结果转换成由 POS, NEG, ZTR, IZ 组成的码流,若为 POS,则系数恢复成系数的值加上 $3/2 K$,若为 NEG,则系数恢复成系数的值减去 $3/2 K$ 。最后一个系数被确定后,将阈值减半,重复上面的过程,就能重构小波系数。当阈值为 1 时,能无失真地恢复小波系数。

5.3 算术编码

算术编码属于熵编码,它为除去数据间的冗余提供了一种有效的机制^{[18][19]}。算术编码不仅可以用在图像编码中,也可以用在对其它类型的数据编码。本文主要用来对小波系数编码后的数据进行再编码,以进一步提高图像数据的压缩率。所以,下面就讨论算术编码的基本思想。

在算术编码前,主要的统计编码是游程编码和霍夫曼编码。理论上讲,霍夫曼编码可以得到最佳的编码效果,但实际上却不行。这是由于实际中由于计算机的限制,霍夫曼编码必须用整数个二进制“位 (bit)”来进行编码,这就使理论上计算得到如: 0.33bit、1.1bit 的码长时,都只能用 1bit 来表示,就使其编码性能大

打折扣。

算术编码的出现，解决了这个问题，由此提高了数据压缩的效果。算术编码用一个介于 0~1 之间的小数来表示对符号编码的结果，不论待编码的数据有多少，其编码结果都是一个数，且是介于 0~1 间的小数。下面以一个实际的例子来说明自适应算术编码的基本思想和工作原理。

一、编码过程

首先作以下假设：信息中仅可能存在的符号为 a、b、c，等待编码的信息为：abcbcb。

那么整个编码过程如下：

(1) 编码前，先默认 a、b、c 出现的概率相等，都为 $1/3$

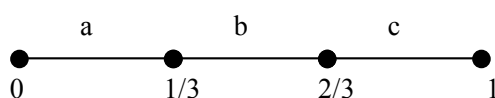


图 5.4 编码前各符号概率

Figure 5.4 Symbol probability before encoding

(2) 当得到信息的第一个字符“a”时，就在上面“a”的区间 $[0, 1/3)$ 中重新划分 3 个符号的概率：

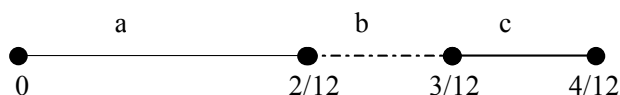


图 5.5 对第一个符号进行编码后的概率

Figure 5.5 The probability after 1st symbol encoding

(2) 得到第二个字符“b”，重复步骤 (2) 的过程：

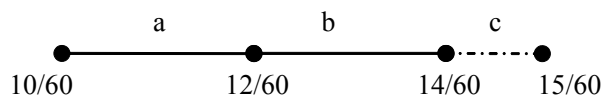


图 5.6 对第二个符号进行编码后的概率

Figure 5.6 The probability after 2nd symbol encoding

(3) 得到第三个字符“c”：

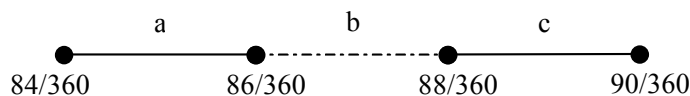


图 5.7 对第三个符号进行编码后的概率

Figure 5.7 The probability after the third symbol encoding

(4) 得到第四个字符“b”:

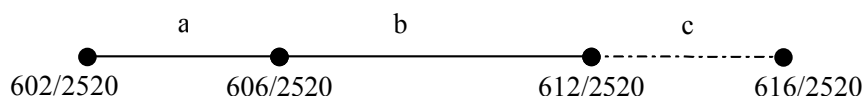


图 5.8 对第四个符号进行编码后的概率

Figure 5.8 The probability after the forth symbol encoding

(5) 得到第五个字符“c”:

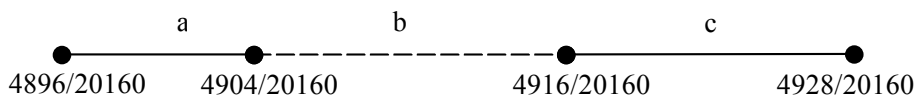


图 5.9 对第五个符号进行编码后的概率

Figure 5.9 The probability after the fifth symbol encoding

待编码的信息段的最后一个字符为“b”，此时就不再需要先前面那样重新划分符号的区间了。而是从最后一次划分区后“b”的区间，即图 5.9 中虚线所示部分—— $[4904/20160, 4916/20160) \Leftrightarrow [0.24325 \dots, 0.24384 \dots)$ ，中选取一个数值输出即可，例如，选择一个中间值：0.2435。到此，整个编码过程结束，编码后的信息即为 0.2435。

二、解码过程

在了解了算术编码方法的编码过程后，再讨论其解码过程就很容易了。

解码过程的步骤同编码时一样，只不过在编码的时候是不断读入待编码的符

号，从而得到不断变化的符号概率区间。而在解码的时候，是利用编码的最终结果 0.2435，通过判断 0.2435 处于当前符号概率的哪个区间，从而得到相应的解码值，然后用这个值又象编码那样来重新划分符号的概率区间。不断重复上述的过程，由判断 0.2435 位于哪个符号的区间内而得到解码值，直到解码完毕，即恢复出了原来的信息。

同编码的时候相比，在什么时候结束解码就需要通过特殊的处理来解决，例如：可以加入一个原符号的长度信息，由此控制解码的结束。总之，解决这个问题并不困难，有多种方法，这里只是指出这个区别引起注意。

以上介绍的算术编码的例子实际上是 0 阶自适应算术编码。它只是利用了符号在已输入信息中的概率，而没有考虑任何上下文信息。因此其编码效率还可以提高。

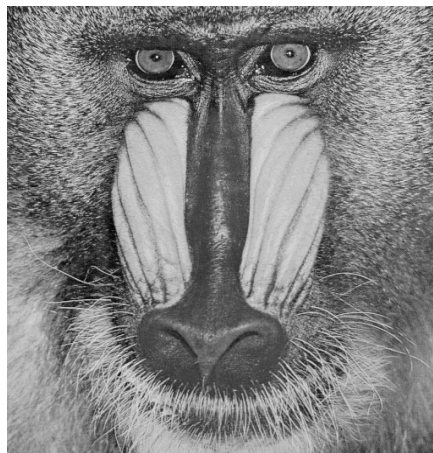
如果将上面的模型改为统计符号在某个特定符号后出现的概率，那么模型就成为了 1 阶自适应模型，同理，如果是统计符号在特定的 n 个符号后的概率，模型就是 n 阶自适应模型。但是随着模型阶数的提高，消耗的系统空间和时间的也随之增加，以目前的技术而言将会难以接受。所以目前使用最多的是 3 阶以下的自适应算术编码器。本文的自适应算术编码采用了文献[21]中的 Codec2003 编码方案。并实现了 2 阶自适应算术编码器。

5.4 测试图像的选取

在数据压缩的研究领域内，都是用相同的测试数据来对不同的算法、方法进行测试，从而比较它们在压缩效率和速度上的性能。在图像压缩领域也有通用的标准测试数据——标准静态图像，图 5.10 就给出了 3 幅最常用的标准图像。



(a) Lena



(b) Baboon



图 5.10 标准测试图像

Fig 5.10 standard test images

图 5.10 (a) “Lena” (或是 Lenna)图像长久以来被广泛地使用在图像压缩成果测试上,研究图像压缩理论的学者们都以这张图像做为样本来比较他们研究成果的优异。同时, Lena 图像不仅是图像压缩界,也是整个图像处理领域最常用到的一幅测试图像。这里顺便介绍一下 Lena 图像的来历。该图是一副数码图像,来源是花花公子杂志 1972 年 11 月号的一张折页海报。是瑞典人 Lena Soderberg 的特写(花花公子杂志中的拼法是 Lenna,Lena 则是瑞典原名的拼法。),被南加利福尼亚大学一位不知名的学者发现后加以剪裁和扫描,用来作为其图像压缩研究的测试图像。从此 “Lena” 在图像处理研究领域广为使用,成为最重要、知名和常用的测试图像。而 Lena 本人一直生活在瑞典,直到 1988 年她才知道自己的 “名声”,1997 年 5 月她还应邀参加了波士顿第 50 届图像科学与技术协会年会。

是什么原因使得 lena 等图像成为检验图像压缩算法的测试图像呢?因为这些测试图像都是很有特点和代表性的。先来分析一下图 5.10 (a) 和 (b) 两幅图像的特点。不难看出,“Lena” 图像多为纯连续色调,特别是墙和露出的皮肤区域;图像中帽子是好的连续色调,而头发和帽子上的羽毛却是差的连续色调;墙上的直线和镜子的曲线部分则是离散色调图像的特征。因此用该图作为图像压缩的测试图是很具有代表性的。

图 5.10 (b) “Baboon”,也称为 “Mandri” 是一幅自然图像,却是一幅差的连续色调图像的典型例子。如果观察该图像的细节,会发现许多像素与其相邻像素差别甚大,这是因为在动物的脸部存在大量的毛发。用任何的方法压缩这幅图像的效果都很差,所以作为测试图像也是相当有代表性的。(c) 也是人物,但背景又比 (a) 有了更多的变化。总之,这 3 幅图像都是典型的测试图像,通过对它们进行处理,容易比较不同压缩算法的性能。

为了便于分析和讨论，以及提高对算法分析的普遍实用性，本文的算法就主要用这 3 幅图像作为测试图像，考虑到简化分析的步骤，便于把注意力集中在算法分析中，以及彩色图像本身可以分解为三个分量图像，基本上可以按照灰度图像的处理进行，所以本文就只采用了灰度图像进行分析和讨论。3 副图像均为 $512 \times 512 \times 8\text{bpp}$ 的灰度图像。图像的像素值大小介于 0 到 255 之间。

6 实验结果及分析

本文分别做了无损压缩和有损压缩的编码实验。所选用的图像是第五章图 5.5 所示的三副标准测试图像。小波选用 5/3 小波和 9/7M 小波，编码算法选用 EZW 算法和本文提出的不完全扫描法。实验是在 Windows2000 操作系统、PIII555MHz、内存 256M、Visual C++6.0 编译环境下进行。

6.1 无损压缩

(1) 5/3 小波

表 6.1 5/3 小波无损压缩比较 (单位: bpp)

Table 6.1 Comparison for 5/3 wavelet lossless compression (bpp)

Bit rate (bpp)	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
2 次变换	5.4627	5.3114	6.5402	6.2117	5.9421	5.8607
3 次变换	5.3596	5.2023	6.4989	6.1322	5.8676	5.7824
4 次变换	5.3342	5.1833	6.4667	6.1206	5.8549	5.7701
5 次变换	5.3303	5.1786	6.4634	6.1194	5.8521	5.6680

(2) 9/7-M 小波

表 6.2 9/7-M 小波无损压缩比较 (单位: bpp)

Table 6.2 Comparison for 9/7-M wavelet lossless compression (bpp)

Bit rate (bpp)	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
2 次变换	5.4103	5.2596	6.5021	6.1357	5.8510	5.7712
3 次变换	5.3152	5.1635	6.4674	6.1006	5.7764	5.6788
4 次变换	5.2986	5.1477	6.4583	6.0913	5.7521	5.6533
5 次变换	5.2967	5.1456	6.4564	6.0894	5.7498	5.6510

从实验结果可见，在无损压缩时，本文算法优于 EZW 算法，以 Lena 图像用 9/7M 小波四次变换为例，压缩比约提高 1.9%。

为了方便比较，把经小波变换后的系数编码效果用图形表示出来，这里选择的数据是对“Lena 图像”作 9/7-M 小波变换后编码的数据，其它图形的情况类似，就不一一绘出。图 6.1 中纵坐标为码率 (bpp)，横坐标为小波变换的次数。

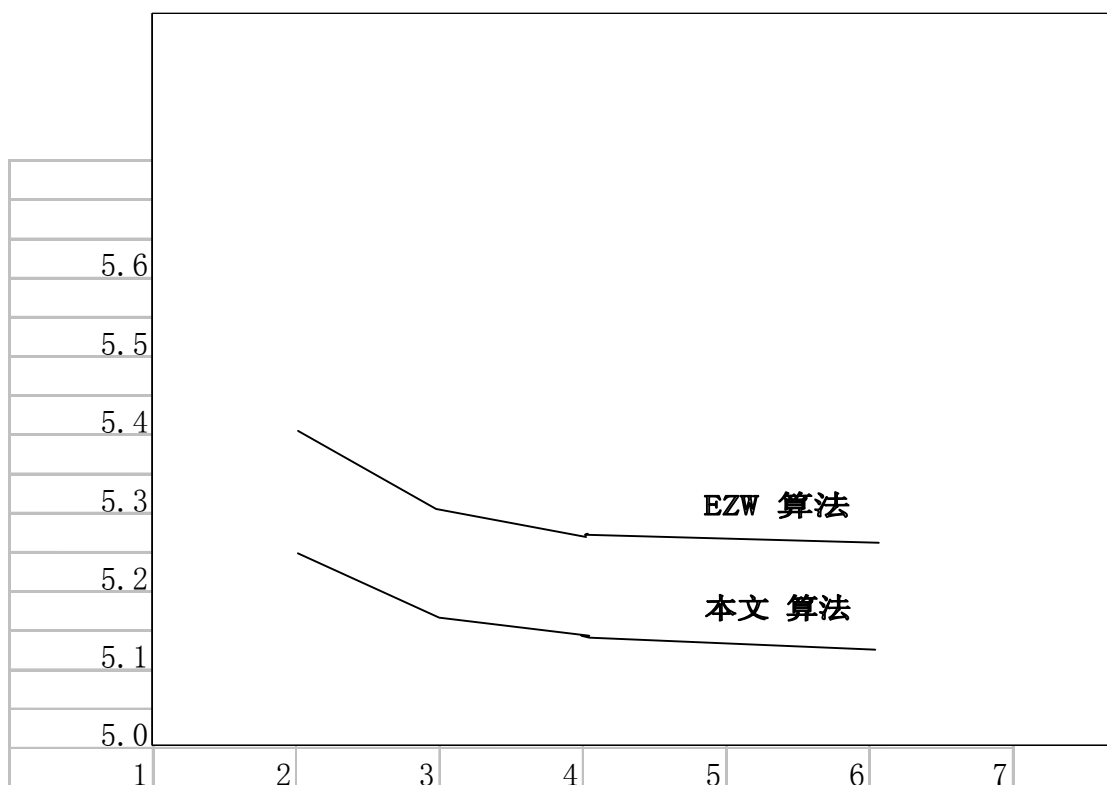


图 6.1 两种算法的压缩效果比较图—Lena

Figure 6.1 comparison for effect of tow algorithms – Lena

从上图中可以更清楚的比较前面表中的数据，可以看出本文的算法优于 EZW 算法。

6.2 有损压缩

(1) 5/3 小波

下面是采用 5/3 小波做四次变换所得结果。

表 6.3 有损压缩时的峰值信噪比 (单位 dB)

Table 6.3 PSNR results of lossy compression (dB)

Bit rate (bpp)	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
1.0	39.55	40.38	32.61	33.32	35.14	36.06
0.5	36.28	36.13	28.10	27.96	30.53	30.27
0.25	33.17	32.24	25.23	23.68	26.77	25.48

(2) 9/7-M 小波

下面是采用 9/7-M 小波做四次变换所得结果。

表 6.4 9/7-M 小波有损压缩时的峰值信噪比 (单位 dB)

Table 6.4 9/7-M wavelet PSNR results of lossy compression (dB)

Bit rate (bpp)	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
1.0	39.62	40.46	32.71	33.46	35.46	36.24
0.5	36.32	36.15	28.19	28.03	30.56	30.33
0.25	33.17	32.21	25.27	22.98	26.83	25.49

从实验结果可见,在有损压缩时两种算法各有优劣,在压缩倍数小于 16 倍时,本文算法略优,高于 16 倍压缩时,EZW 算法略优。因为 EZW 算法对重要系数进行了更精细的级间编码。

为了衡量压缩编码图像的主观质量,图 6.2 展示了 Lena 图像采用本文算法经四次小波变换在不同压缩比下编码后重建图像的结果。



图 6.2 本文算法对 lena 图像的编码结果

Fig 6.2 Performance of proposed method operation on lena

由上图可知，在压缩比为 8: 1 时基本上和原图无差别；在压缩比为 16: 1 时，视觉效果和原图有稍微差别，视觉效果较好，细节丰富；在压缩比 32: 1 时视觉效果有一定的失真。我们注意到在不同压缩比下的图像都没有方块效应，这也正是小波变换优于 DCT 变换的地方。

6.3 编/解码时间

这里比较了编码算法在无损压缩和有损压缩情况下的编/解码过程执行时间。下面以 4 次 5/3 小波变换为例。

(1) 表 6.5 和表 6.6 分别对 EZW 算法和本文算法在无损压缩时的编/解码时间进行了比较。

表 6.5 图像无损压缩的编码时间比较 (单位为秒)

Table 6.5 Process time of lossless encoding (s)

	Lena	Baboon	Barbara
EZW	13.32	11.068	13.820
本文算法	11.463	9.254	11.352

表 6.6 图像无损压缩的解码时间比较 (单位为秒)

Table 6.6 Process time of lossless decoding (s)

	Lena	Baboon	Barbara
EZW	10.249	8.330	10.485
本文算法	10.946	8.803	10.896

(3) 表 6.7 和表 6.8 分别对 EZW 算法和本文算法在有损压缩时的编/解码时间进行了比较。

表 6.7 图像有损压缩的编码时间比较 (单位为秒)

Table 6.7 Process time of lossy encoding (s)

Bit rate	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
1.0(bpp)	6.342	5.812	5.321	4.810	6.393	5.823
0.5(bpp)	5.879	5.401	4.872	4.509	5.891	5.452
0.25(bpp)	5.740	5.289	4.730	4.370	5.779	5.291

表 6.8 图像有损压缩的解码时间比较 (单位为秒)

Table 6.8 Process time of lossy decoding (s)

Bit rate	Lena		Baboon		Barbara	
	EZW	本文算法	EZW	本文算法	EZW	本文算法
1.0(bpp)	5.732	5.879	4.631	4.710	5.723	5.769
0.5(bpp)	5.152	5.249	3.971	4.112	5.146	5.231
0.25(bpp)	4.933	4.991	3.792	3.810	4.932	4.959

通过对编码和解码时间进行统计, 在无损压缩时本文算法平均编码时间只相当于 EZW 算法编码时间的 84%, 本文算法明显优于 EZW 算法。解码时间 EZW 算法比本文算法要快 5.2%, EZW 算法略优。

在 EZW 算法中所有重要系数都要在附表中保存, 直到编码结束。本文算法不需要对重要系数进行保存, 从而减少了对内存的开销, 这是本文算法优于 EZW 算法的另一个重要方面。表 6.9 是三幅测试图像小波变换后非零系数的统计结果, 反映了本文算法节省内存的大小。如 Lena 图像编码时, 若每个重要系数用 1 个字节表示, 保存所有的重要系数需要占用内存 223780 字节。

表 6.9 非零系数的统计

Table 6.9 The number of nonzero coefficient

测试图像	Lena	Baboon	Barbara
非零系数的个数	223780	233565	215560

7 全文总结与展望

7.1 主要结论

随着社会的进步和多媒体技术的日益成熟,对图像通信的需求也越来越多。阻碍图像通信应用的瓶颈在于图像信号数据量巨大,必须采用有效的压缩手段对其进行压缩处理才能满足大多数系统存储、传输的要求。小波变换因其特有的与人眼视觉特性相符的多分辨率分析及方向选择能力,而被广泛应用于图像编码领域,取得了很大的成功。近年来,提升方法的提出为小波的构造提供了新思路,也为基于小波变换的图像编码注入了新的活力。

本文对基于小波的图像压缩编码方法进行了研究。主要取得了以下研究成果:

1. 对提升算法进行了研究,采用提升算法实现小波变换,相对于传统小波变换,提升算法处理速度快,占用资源少,可实现整数小波变换。
2. 对小波编码算法进行了研究,提出了一种改进的零树小波编码算法---不完全扫描法。对变换后的小波系数选用零树编码算法,考虑到实时应用和实际应用的环,对经典的 EZW 算法进行了改进,提出了不完全扫描法,它相对于 EZW 算法具有系统开销小,操作简单的优点,提高了编码速度。
3. 最后,本文用 C/C++语言编写了图像压缩编解码器。该编解码器能静止图像进行不同比特率的压缩编码。通过对标准测试图像的压缩编码实验,证实本文的图像压缩方法能有效进行图像压缩,提高了编码速度,能实现任意比特率压缩,编码输出嵌入式码流。

7.2 后续研究工作的展望

通过一年多的研究工作,对于小波变换、零树编码、提升算法以及熵编码技术有了一定的理解和掌握,也进行了一些分析,设计与仿真。可以说这样的研究方向是很有发展前景的,在已有的研究基础上朝着这一个方向前进,可以对未来的研究工作做一个很好的规划。今后可向以下几个方面进一步研究和开发:

1. 优化小波变换和逆变换,使其处理的速度更快,符合实时应用的要求。
2. 对零树编码中的搜索路径进行优化,提高其搜索速度。
3. 加入各种先进的编码模式,近一步提高其编码效率。
4. 将目前完成的 C\C++语言程序转换成汇编语言,移植到 DSP 中及设计相关的辅助硬件电路部分。

总之,基于小波变换的图像压缩编码是一个新兴的领域,出现了许多新的优

秀压缩算法。是与多学科有联系的交叉学科，其研究范围十分广泛，在图像压缩中的应用的涉及面也很广，其中的许多理论和实际问题还有待进一步解决，技术发展十分迅速，新的挑战层出不穷。本文作者一直从事这方面的学习和研究工作，并在本文的撰写过程中付出了辛勤的汗水，但由于图像压缩技术的发展日新月异，而且本文作者从事这方面工作的时间不长，在理论、实践方面的知识和经验毕竟有限，文中难免存在疏漏和不足之处，真诚地欢迎广大专家和读者予以批评指正。

致 谢

本文的研究工作和毕业论文的撰写是在我的导师朱冰莲副教授的精心指导和悉心关怀下完成的。在整个论文工作中，不仅为论文的选题、试验方案的确定倾注了大量的心血，在繁忙的工作中多次抽出时间与我进行探讨，而且为我平时的学习和工作提供了许多便利条件。朱老师严谨扎实的学者作风，诲人不倦的师长风范，对待科研一丝不苟的工作态度使我深受影响，导师的言传身教将会使我受益终生！在此表示最诚挚的谢意。

感谢重庆大学通信工程学院的全体老师，三年来给予我诚挚的关怀和帮助，并引导我在学习和生活中找到正确的方向，我会铭记他们的谆谆教诲。

感谢与我同一实验室的孔杰、苏璞、孔召贤、梁立宏、程联营、张文明、吕佶、陈禧等同学，与他们的友好交往、有益的学术讨论使我学到了很多，我与他们共同度过了一段难忘的学习生活。

我还要感谢我的父母、女友，正是他们一如既往地给予我生活上、精神上的支持，使我顺利完成了整个研究生阶段的学习。

衷心地感谢在百忙之中评阅论文和参加答辩的各位专家、教授！

王文涛

二〇〇五年五月 于重庆

参考文献

- [1] 崔锦泰〔著〕, 程正兴〔译〕.小波分析导论.西安交通大学出版社, 1995
- [2] 程正兴.小波分析算法与应用.西安交通大学出版社, 1998
- [3] M .Unser. Approximation power of biorthogonal wavelet expansions. IEEE Trans.on Signal Processing, 1996, 4(4):519-527
- [4] M .Natonini. Image coding using wavelet transform. IEEE TranS. on Image Processing, 1992, 1:205-220
- [5] J .Villasensor. Wavelet filter evaluation for image compression . IEEE Trans.on Image Processing, 1995, 4(8):1053-1060
- [6] Li H ,Wang , Wu L,A novel design of lifting scheme from general wavelet. IEEE Trans. Signal Processing, Vol.49,pp:1714-1717,2001
- [7] Cheng L Z, Xiang D L. Popular biorthogonal wavelet filters via a lifting scheme and its application in image compression, IEEE Proceedings: Vision, Image and Signal Processing, Vol.150(4),PP: 227-232,2003
- [8] O.Rioul, P.Duhamel. Fast algorithms for discrete and continuous wavelet transforms . IEEE Trans. Information Theory.1992, 38(2):569-585
- [9] Sweldens W.The lifting scheme:a construction of second generation wavelets [technical report].USA: University of South Carolina, 1995
- [10] Sweldens W. The lifting scheme:a custom — design construction of biorthogonal wavelets. APPL. Comput. Harmon. Anal., 1996, 3(2):186-200
- [11] Honeying Meng, Zhihua Wang. Fast spatial combinative lifting algorithm of wavelet transform using 9/7 filter for image block compression . Electronics Letters, 2000, 36(21)
- [12] A R Calderbank, I Daubechies .Wavelet transforms that map integers to integers. Applied and Computational Harmonic Analysis, 1998, 5(3):332-369
- [13] Cheng L Z, Zhong G J, Luo J S.A new family of integer lapped transform that can map integers to integers . IEE Proceedings: Vision, Image and Signal Processing, Vol.149(2),pp:91-96,2002
- [14] Reichel J, Menegaz G, Nadenau M J, et al. Integer wavelet transform for embedded lossy to lossless image compression, IEEE Trans. Image Processing, Vol.10,pp: 383-392, 2001
- [15] M .D.Adams, F.Kossentini. Reversible integer-to-integer wavelet transform for image compression: Performance evaluation and analysis.IEEE Trans. Image Processing, 2000, 9(6):1010-1024.

- [16] J .Reichel, M.Menegaz. Integer wavelet transform for embedded lossy to lossless image compression. IEEE Trans. Image processing, 2001, 10(3):383-392
- [17] Shapiro J M.Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans. on Signal Processing, 1993, 41(12):3445-3463
- [18] A .Said, W.A.Pearlman.A new, fast, and efficient image codec based on set Partitioning in hierarchical trees. IEEE Trans. Circuits and Systems for Video Technolog,1996, 6(3):243-250
- [19] A .Islam, W.A.Pearlman.An embedded and efficient low-complexity hierarchical image codec. Proceedings SPIE 3653 , Visual Communications and Image Processing,99, San Jose, CA.
- [20] Park K H, Park H W. Region-of-interests coding based on set partitioning in hierachical trees. IEEE Trans. On CSVT, 2002, 12(2):106-113
- [21] 苏冬, 基于整数小波的图像压缩编码方法, [重庆大学学位论文].重庆: 重庆大学, 2004.5
- [22] 于景侠, 陈雷霆.基于小波变换的静止图像编解码算法研究与实现,[电子科技大学学位论文].成都: 电子科技大学, 2003.5
- [23] 冯岩, 小波变换在图像压缩中的应用, [华中师范大学学位论文], 武汉: 华中师范大学, 2004.5
- [24] 桂乐, 小波分析理论及其在图像压缩中的应用, [西北大学学位论文], 西安: 西北工业大学, 2003.4
- [25] 黄贤武、王加俊、李家华, 数字图像处理与压缩编码技术, 电子科技大学出版社, 2000.12.
- [26] Paul G.Howard and J.S. Vitter . Arithmetic coding for data compression. Proceedings of the IEEE, 82(6):857-864. 1994.6.
- [27] I.H. Witten , R.M. Neal , J.G. Cleary. Arithmetic coding for data compression . Communication of the ACM, 30(6): 520-540. 1987.6.
- [28] 丁绪星、朱日宏、李建欣, JPEG2000 及其新技术, 计算机应用, 2003, (2) .
- [29] ISO/IEC.JCTI/SC29 CD109180I:JPEG:Digital compression and coding of continuous tone still image, 1992
- [30] Christopoulos C, Askelof J, et al .The JPEG 2000 Still Image Coding System: An overview. IEEE Transaction on Consumer Electronics, 2000,46(4)
- [31] Christopoulos C, Askelof J and Larsson M. Efficient encoding and reconstruction of Regions of Interest in JPEG 2000.European Signal processing Conference, Tampere, Finland,2000
- [32] ISO/IEC 15444-1: Information technology—JPEG 2000 image coding system—Part1 : Core coding system,2000
- [33] Aizawa K, et al. Model-based analysis synthesis image coding system. Signal Processing: Image Communication, 1989,(1):139-152

- [34] ITU-T Recommendation H.263, "Video coding for low bit rate communication", 1996
- [35] ITU-T. Draft for "H.263++" annexes U, V, and W to recommendation H.263, 2000
- [36] ISO/IEC.JCT1/SC29 CD11172-2 MPEG1, international standard for coding of moving pictures and associated audio for digital storage media at up to about 1.5Mb/s. 1991
- [37] ISO/IEC.JCT1/SC29 CD13818-2 MPEG2, coding of moving pictures and associated audio. 1993
- [38] ISO/IEC.JCT1/SC29 WG11/N3536 MPEG-4:overview V.15.Beijing, 2000
- [39] K.R.Castleman, 朱志刚等译, 数字图像处理, 电子工业出版社, 1998,
- [40] Sweldens W. The lifting scheme: a custom-design construction for biorthogonal wavelets. Technical Report, University of South Carolina, 1994:1-39
- [41] Wim Sweldens, The Lifting Scheme: A Construction of Second Generation Wavelets, 1995
- [42] Ingrid Daubechies, Wim Sweldens. Factoring Wavelet Transforms into Lifting Steps. Technical report, Bell Laboratories, Lucent Technologies. 1996.
- [43] W.Sweldens, The lifting scheme: A construction of second generation wavelets, SIAM J.Math. Anal.29(2)(1997),511-546.
- [44] 周宁, 汤晓军等. 提升小波快速算法及其在 JPEG2000 中的应用. 中国有线电视, 2002 (18) 6-10
- [45] Michael D.Adams, Reversible Integer-to-Integer Wavelet transforms for image compression: IEEE Trans. On CSVT, 2001, 10(2):96-100
- [46] 张荣, 俞能海, 刘政凯, 吴文忠. 基于整数小波变换的多光谱图像无损压缩. 遥感学报.Vol.5 No.1. 2001.
- [47] 王相海, 张福炎. 一种基于零树小波的图像比率可分级编码方法的研究. 南京大学学报 (自然科学). Vol.38 No.2. 2002.
- [48] Asad Islam, William A.Pearlman. An Embedded and Efficient Low-Complexity Hierarchical Image Coder. Proceedings SPIE 3653, Visual Communications and Image Processing '99, San Jose, CA, Enero 1999.
- [49] Ulug Bayazit, William A.Pearlman. Algorithmic Modifications to SPIHT. IEEE Int. Conf. on Image Processing (ICIP 2001), Thessaloniki, Greece. Oct. 2001.
- [50] David Salomon 著. 吴乐南 等译. 数据压缩原理与应用. 第二版. 电子工业出版社. 2003.9
- [51] 王国秋, 袁卫卫. 一般的 9-7 小波滤波器及其图像压缩性能研究. 电子学报, 2001, 29 (1)
- [52] Cheng L Z, Xiang D L, Zhang Z. General construction of 9/7 wavelet filter and its application in image coding. Optical Engineering, Vol.42(7), pp: 1949-1955, 2003
- [53] Cheng L Z, Zhang Z H, Xu H. General construction of 9/7 wavelet filter and its application in

- image compression . ICSP, Aug, 2002,pp: 355-359, Beijing
- [54] Michael D.Adams. Reversible Integer-to-Integer Wavelet Transform for Image Compression[Dissertation]. The university of British Columbia. 2002.9. pp.86~108.
- [55] 彭进业, 俞卞章, 张烨, 刘宏兵. 一种改进的图像自适应零树编码方法. 电子学报. Vol.30 No.4. 2002.
- [56] 杨云志, 黄顺吉, 王建国. 自适应量化编码的研究. 电子技术, 2001 年第一期 374~386
- [57] 朱秀昌, 刘峰, 胡栋. 数字图像处理与图像通信, 北京邮电大学出版社, 2002.5, 第 1 版, 汇源. 数字图像通信原理与技术. 国防工业出版社

附 录

作者在攻读硕士学位期间参加科研

- [1] 2004.9~今 作为课题组主要成员，参加研究生创新实验基地课题“OFDM 自适应调制技术研究”。

作者在攻读硕士学位期间发表的论文

- [1] 朱冰莲，王文涛. “基于不完全扫描法的彩色图像压缩方法”. 重庆大学学报已录用

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得重庆大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

签字日期：2005 年 4 月 25 日

学位论文版权使用授权书


本学位论文作者完全了解重庆大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权重庆大学可以将学位论文的全部或部分内 容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

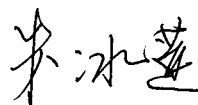
保密（☐），在____年解密后适用本授权书。

本学位论文属于

不保密（☐）。

（请只在上述一个括号内打“√”）

学位论文作者签名：

导师签名：

签字日期：2005 年 4 月 25 日

签字日期：2005 年 4 月 25 日