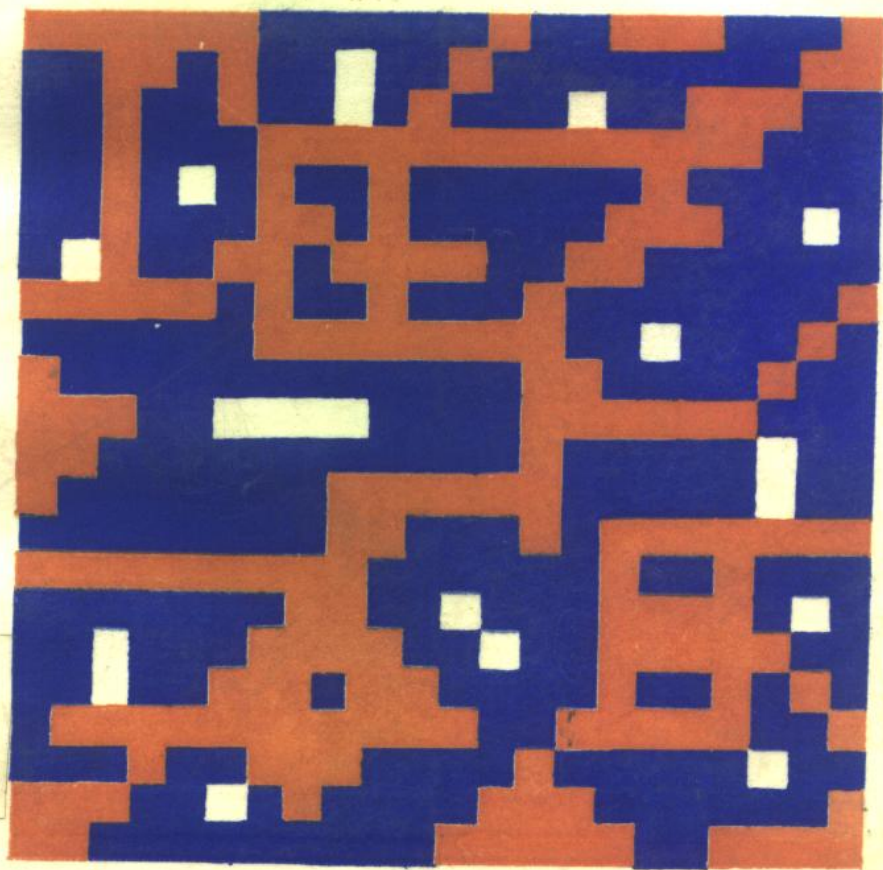


数据图像压缩编码

冯玉琨 邵玉明 张 星 编译
袁保宗 审校



中国铁道出版社

数据图像压缩编码

冯玉珉 邵玉明 张 星 编译
袁保宗 审校

中国铁道出版社
1993年·北京

(京)新登字 063 号

内 容 简 介

本书针对图像存贮与通信中的关键问题,在简单回顾了有关信源熵和率失真理论之后,综述了无损和有损图像压缩的编码技术分类、基本原理和性能比较;然后具体介绍了递归块图像数据压缩方面的几种编码技术、算法,并举例给出实验数值结果及目视效果图片。全书内容新颖,有实用价值。

本书分为两个层次:前三章适于一般信息工作者通晓图像编码技术参考;后七章更适于从事图像处理与图像通信的科技人员参考,也可供大专院校师生和研究生作为学习编码技术与熟悉算法的参考书。

数据图像压缩编码

冯玉琅 邵玉明 张 康 编译
袁保宗 审校

*

中国铁道出版社出版、发行

(北京市东单三条 14 号)

责任编辑 黄成士 封面设计 翟 达

各地新华书店经售

北京市燕山联营印刷厂印

开本: 850×1168 毫米 1/32 印张: 12 字数: 272 千

1993 年 12 月 第 1 版 第 1 次印刷

印数: 1—1500 册

ISBN 7-113-01476-3/TP·147 定价: 10.60 元

前 言

人类进入信息时代，信息化、高技术化成为当今世界新技术革命的重要特征，并正在推动现代科学技术和社会生产力的飞跃发展，加速着人类物质文明与精神文明的步伐。世界各个活动领域中日益丰富与活跃的信息与信息技术，使社会生活展现出绚丽多彩的新局面。近十年来，计算机技术与通信有力结合的新成就，不断满足着人类逐日增长的信息需求。传统的信息业已被突破，电话通信现代化就是其显著标志之一；但人类获取信息总量中，听觉信息只占 20%，而视觉信息却高达 60~70%，所以人类对于形象、生动、真实、全面的图像通信具有更大需求。

由通信的广义定义，就图像信息而言，图像通信就是图像信息的时/空转移，即广义图像通信要包括图像的存取与传输。为了在有限的存贮空间和限定的传输带宽条件下，有效地存贮与传送图像信息，通常采用有效的压缩编码技术，以实现用尽可能低的编码比特率而满足一定的重建图像质量的需求。

近些年来关于图像处理的新书已有多本，图像通信的科普读物及系统性教材、参考书国内已有出版；但专为图像信息时/空转移而编写的有关压缩编码技术书目，尚不多见。作者在参考了几本近年来出版的外文专著和 IEEE 等刊物有关研究论文后，编写了这本小书。其中后半部主要是编译内容，以期为读者尽快提供出有关压缩技术的新颖内容。

本书首先介绍了以信源熵和率失真理论为主要内容的信息论基础知识，它为压缩提出了理论限界。随后介绍了无损与有损压缩的各种技术原理及性能。本书后面的各章节主要介绍几种压缩编码算法，并举例给出了数值结果与目视图片。这些章节内容统

归较新的递归块图像数据压缩编码，具有较高的参考价值。

本书由冯玉珉主编，并编写 1、2、3 章和全书构思编排；邵玉明编译 4、6、8、10 各章，并进行了全书文字校核工作；张星编译 5、7、9 章及附录。全书由袁保宗教授审校。

由于本书编写仓促，作者从事图像通信研究阅历不长，书中必有不当或错误之处，敬请读者批评指正，作者将不胜感谢。

编 者

于北方交通大学

1992. 10.

目 录

1	数字图像和图像压缩概述	(1)
2	信息论基础	(9)
2.1	信源模型和熵	(9)
2.2	变长编码	(16)
2.3	熵估值和无损压缩	(26)
2.4	率失真理论和有损压缩	(32)
3	关于无损与有损图像编码	(34)
3.1	引 言	(34)
3.2	位面编码	(35)
3.3	无损预测编码	(39)
3.4	有损加无损残留编码	(43)
3.5	有损预测编码	(44)
3.6	变换编码	(51)
3.7	块截短编码	(57)
3.8	矢量量化	(60)
3.9	子带编码	(63)
3.10	分层编码	(66)
3.11	关于有损压缩技术的选择	(72)
4	递归块编码——算法引出的理论	(79)
4.1	前 言	(79)
4.2	建 模	(79)
4.3	自回归模型	(80)
4.4	非因果模型	(84)
4.5	二源分解	(86)
4.6	一维 RBC 算法	(89)

4.7	一阶 AR 模型的边界响应	(92)
4.8	二维最小方差模型	(100)
4.9	二维非因果图像模型实例	(101)
4.10	通过变换方法的二维边界响应	(104)
4.11	二维 RBC 算法	(105)
4.12	近似的边界响应	(107)
4.13	RBC 的优点	(122)
5	变换编码的比特分配和量化	(124)
5.1	引 言	(124)
5.2	变换与量化器的选择	(125)
5.3	比特分配	(128)
5.4	整数比特分配	(136)
5.5	零级量化器	(139)
5.6	比特分配的 RD 函数选择	(141)
5.7	RBC 率失真分析	(144)
5.8	DCT 与 RBC 综合设计	(157)
5.9	彩色编码	(165)
6	区域编码	(167)
6.1	引 言	(167)
6.2	原始图像	(168)
6.3	仿真设施	(173)
6.4	图像质量估测	(175)
6.5	一维编码效果	(182)
6.6	二维编码结果	(199)
6.7	混合编码	(216)
6.8	结 论	(234)
7	活动图像自适应编码	(236)
7.1	引 言	(236)
7.2	自适应 DCT	(237)

7.3	自适应RBC	(250)
7.4	编码结果	(259)
7.5	结 论	(270)
8	基于四叉树分割的自适应编码	(275)
8.1	引 言	(275)
8.2	自适应分割	(275)
8.3	四叉树方法	(276)
8.4	分割的编码	(282)
8.5	重 建	(288)
8.6	残留编码	(296)
8.7	结 论	(308)
9	四叉树残留的矢量量化 (QVQ)	(312)
9.1	引 言	(312)
9.2	矢量量化	(312)
9.3	差值和内插VQ	(321)
9.4	四叉树残留的VQ	(322)
9.5	编码结果	(325)
9.6	结 论	(327)
10	总 结	(330)
10.1	引 言	(330)
10.2	新 成 果	(330)
10.3	编码建议	(332)
10.4	今后的工作	(334)
附录A	常微分、偏微分和差分方程	(337)
A. 1	引 言	(337)
A. 2	二阶常微分方程	(337)
A. 3	二阶常差分方程	(338)
A. 4	二阶偏微分方程	(340)
A. 5	二阶偏差分方程	(342)

附录 B 离散正弦变换特性	(345)
B. 1 引言	(345)
B. 2 DST 计算技术	(346)
B. 3 指数序列	(346)
B. 4 常数序列	(347)
B. 5 线性序列	(347)
B. 6 双曲序列	(348)
B. 7 正弦序列	(349)
附录 C 变换域方差分布	(350)
C. 1 引言	(350)
C. 2 一维 RBC	(350)
C. 3 一维 DCT	(353)
附录 D 基于活动类自适应编码的编码参数	(355)
D. 1 引言	(355)
D. 2 自适应 DCT 编码参数	(355)
D. 3 自适应 RBC 编码参数	(366)
参考文献	(375)

1 数字图像和图像压缩概述

近几年来,对以数字形式处理图像的需求显著增加。由于图像扫描器性能的改善以及成本大大降低,相片、印刷文本和其它介质现在可以很容易地转换成数字形式(图像数字化)。数字图像(景像数字化)的直接获取也像传感器以及相关的电子学发展一样更加普遍;在地球遥感技术中例如 LANDSAT 卫星拍照的利用,以及消费市场中电子静物摄像机的出现就是很好的例证。此外,医学中很多不同的摄像方式,诸如计算层析 X 射线摄影法(CT)或磁共振摄像(MRI),直接以数字形式产生图像。计算机生成图像(合成图像)也成为数字数据的主要来源。在广告和文娱活动中广泛应用计算机制图,并且它在科学显像以及工程应用中也在迅速增长。对数字图像如此感兴趣的原因十分明显:用数字形式表示图像可使可视信息以有效、新颖方式灵活地加以控制。在过去的十年中,随着计算能力按指数规律增长的事实,致使数字图像系统在以下各个领域广泛应用,如天文学、遥感、医学、摄影报道、图形艺术、法律实施、广告和制造业等。

尽管数字图像存在很多优点,但有一潜在的问题,就是要表示它们则需要大量比特数。幸而,在数字图像的标准表示中,通常含有大量冗余度。图像压缩是图像数据有效编码的技术/科学,其目的在于获得这种冗余优点,从而减少表示图像所需的比特数。这样就使存贮图像所需的存贮器,或者传输图像所需的通道容量大大节省。近几年来,IEEE 等国际著名刊物发表了不少有关图像压缩的文章;还有几种有关图像编码的专刊;国内外陆续出版了不少有关图像处理与通信的专著、教科书或参考书,有的书目比较详细地介绍了图像存贮与通信压缩技术。

1.1 数字图像的形成

数字图像源可以是现实世界的三维景象，或者可以是预先形成的二维图像，例如相片。无论哪种情况，其源信息都是在空间和幅度上连续的模拟形式。为了产生数字图像，首先利用某种传感器（如果是彩色或多谱数据，则需要多个传感器），在离散位置上对图像源进行取样。这些样值称为像素（作为图像元素）。由传感器产生的这些像素值在某些有限范围是连续的，并且它们与每个取样位置的辐射能强度通常呈线性关系。由于传感器对超出人的视觉范围的波长十分敏感，故考虑限于利用辐射能（而不是光）。在预先已产生图像（诸如相片或幻灯片）情况下，其辐射能是照射到该图像的光源，反射或传送到传感器的结果。

对于这些取样单元可能采取不同的策略，但最常用的是等间隔方格。理论上，每个样本对应于一个无穷小的源图像方格，但由于具体传感器及相关的镜片性能的原因，实际上它是某些有限区域的积分。每单位区域的取样单元数确定系统的取样速率，并且应根据奈奎斯特（Nyquist）取样定理选择其取样速率，即取样速率至少应为图像源最高空间频率分量的2倍。如果不满足这一准则，当图像的空间频率高于取样速率的一半时，经取样的图像信号将出现混迭谱。为了防止混迭失真，系统在取样之前可利用前置滤波器，以限制图像源的带宽。（注意：由传感器完成的积分也起到滤波器作用。）在图像数字化系统中，取样速率通常由扫描分辨率给出，扫描分辨率是取样速率的倒数。通常所需的扫描分辨率取决于实际应用。例如，放射学家以14in的典型距离观察14×17in的射线照片，扫描分辨率为70微米/像素，其俘获空间频率高达7周/毫米足以满足要求（基于人工视觉系统（HVS）对比敏感度函数的数据）。对于35mm的底片须经连续放大和处理，扫描分辨率为更典型的12微米/像素。

为了形成数字图像，将每个取样点的连续值量化成离散电平数之一。众所周知，人工视觉系统同样不能察觉亮度值的均匀变化。例如，亮度值的均匀量化（即线性间隔量化），可导致暗区的量化误差显而易见。由于在多数应用中是通过人来观察图像，故在与人工视觉系统的识别一致的信号域中完成量化尤其重要。在众多情况中，量化之前像素值须经非线性处理（诸如对数或立方根函数），它近似于HVS的非线性。足够用以表示一幅图像所需的量化电平数，也视实际应用而定。对于二元文本文件，由于每个取样点不是黑色就是白色，所以只需两个电平（1比特/像素）。对于自然景像或连续色调照片，通常利用8比特/像素（256个电平）。然而，根据图像源的动态范围及传感器输出类型（例如线性或对数），可能需利用10或12比特/像素。事实上，为了克服非线性直接模拟实现的不准确性，更为复杂图像的数字化是首先利用12~14bit，用线性间隔量化像素值，将非线性用于数字数据，然后利用查表重新量化到8bit。

实际的图像采集过程，往往是在获取图像的同时就完成了取样和量化；也有一些情况是先获取模拟图像，然后去考虑外部数字化问题。例如电视工业用的传统摄像管摄像机，产生一维图像信号，是在时域中和垂直方向上进行取样来表示图像的。更为先进的固态电偶设备（CCD）摄像机也可沿水平扫描线对图像进行取样，所以摄像机的输出就是纵横两个方向上的样本。CCD阵列的输出是一组对应于每个像素亮度的模拟幅度。然而摄像机的输出可以是模拟的（RS-170），也可以是数字的，并主要取决于摄像机内部是否含有模-数转换（ADC）设备。

单色（黑-白）图像仅包含一个通道，而彩色图像则有三个通道，每个通道包括彩色坐标系统中一个分量的样本。该系统可能是三基色（R、G、B——红、绿、蓝），或亮度与色度（色差）信号，即YUV或YIQ。卫星图像数据却包含更多的信道，例如LANDSTA（地球资源技术卫星）用7个频段产生图像。若利用与

一般 TV 图像质量相同的速率取样, 那么一帧图像可包含 $512 \times 512 \times 8 = 2.10$ 兆比特/通道。如果考虑彩色图像序列, 比如广播电视, 其帧频为 25Hz (PAL 制) 或 30Hz (NTSC 制), 每个通道以相同的速率进行取样, 且利用 8 比特/通道量化各像素, 数据速率则为 188.74Mbit/s。

当遇有宽带视频源信号, 但只具备窄带通信线路或存贮介质时, 例如计算机网络、电话线路、磁盘或磁带等, 我们就容易理解数据压缩的必要性。具体来讲, 数据压缩广泛应用于单帧图像传输、广播电视传输, 电视会议、远程船舶导航、文献档案系统、数字放射网络、大规模工程绘图系统等所需的图像数据库和传输系统及存取。

通过下面一些例子可以说明图像数据压缩的必要性。

1.2 图像压缩的必要性

当计算由典型取样速率和量化方法引起的每幅图像的比特数时, 图像压缩则显得非常必要。例如, 考虑以下类型图像所需的存贮量:

- (1) 低分辨率, TV 质量, 彩色电视图像: 512×512 像素/色彩, 8 比特/像素, 和 3 色 $\Rightarrow \approx 6 \times 10^6 \text{ bit}$;
- (2) 以 $12\mu\text{m}$ 扫描的 $24 \times 36\text{mm}$ (35mm) 照像底片: 3000×2000 像素/色彩, 8 比特/像素, 3 色 $\Rightarrow \approx 144 \times 10^6 \text{ bit}$;
- (3) 以 $70\mu\text{m}$ 扫描的 $14 \times 17\text{in}$ 放射照片: 5000×6000 像素, 12 比特/像素, $\Rightarrow \approx 360 \times 10^6 \text{ bit}$;
- (4) LANDSAT Thematic 测绘仪景像(用于遥感): 近似 6000×6000 像素/谱带, 8 比特/像素, 6 个非热能谱带 $\Rightarrow \approx 1.7 \times 10^9 \text{ bit}$ 。

很明显, 甚至几幅图像的存贮就可看出问题。需要图像压缩的另一例子是, 考虑在电话线上传输低分辨率 $512 \times 512 \times 8$ 比特

/像素 $\times$ 3 色的电视图像。当利用 9600 波特 (bit/s) 的调制解调器 (modem) 时, 对于单幅图像的传输占用大约 11 分钟, 对于大多数应用来说, 这是不可接受的。

1.3 压缩技术的分类

上节的例子清楚地说明, 数字化过程对每幅图像产生多么大量的比特数。然而, 由于冗余度的原因, 表示图像信息实际所需的比特数明显减少。通常, 数字图像的冗余度有三种表示方法:

- (1) 空间冗余度, 它是由相邻像素值间的相关性所造成;
- (2) 频谱冗余度, 它是由不同彩色平面 (例如, 在红、绿、蓝 (RGB) 彩色图像中) 或频谱带 (例如遥感的航摄照片) 的相关性所造成;
- (3) 时间冗余度, 它是由图像序列不同帧的相关性造成的。

图像压缩研究的目的在于, 通过消除这些冗余度来减少表示图像所需的比特数。此外, 对于所给定类型的图像, 这是利用信息论的概念完成的。除了这些基本目标外, 对于研究适于不同应用的各种算法, 它也是必不可少的。图像压缩有很多方法, 但它们基本上可分成有损和无损两类。

在无损压缩 (也称为比特保存或可逆压缩) 中, 压缩后在逐像素基础上的重建图像, 其数值等同于原图像。显然, 无损压缩没有信息损失, 达到理想要求。然而, 只有适量的压缩才有可能实现。

在有损压缩 (亦称为不可逆压缩) 情况中, 相对于原图像来说, 其重建图像有损耗成分。对于非常高的压缩率, 其结果与无损压缩差别较大。通常, 在以较大失真为代价情况下可得到较高的压缩率。注意这些损耗是不是显而易见是重要的。实际上, 经常利用目视无损这一术语来表征有损压缩方法, 这样可使在一般观察条件下无明显损失。遗憾的是目视无损的定义相当主观, 在

其解释中应取极其慎重的态度。在某些观测条件下,例如以 4ft 距离观测的 19in 视频监视器,其目视无损算法有可能在更严格观测条件下(诸如 14×17in 印在胶片上的再生图像),导致可见损耗。

根据原输入图像的性质,这两类图像可进一步再分。该图像可以是二元的,例如文本和文件;或者为连续色调,例如 8bit 视像,12bit 医学图像等。它可以是含空间冗余度(如果是彩色图像则为频谱冗余度)的静止图像,也可以是含时间冗余度的图像序列,例如活动图像。图像序列的压缩问题不做专门论述,但对静止图像描述的很多技术可应用于序列压缩。

1.4 压缩数字化参数的作用

数字图像中存在的冗余度,在很大程度上取决于形成图像所用的系统,以及表示图像所用的参数。尤其是,取样速率,量化电平数,及存在于图像源和/或传感器的噪声都能影响所实现的压缩。虽然图像间将存在变化差别,但对于统计基础上的压缩算法通常可形成下面的趋势:

- (1) 随着取样速率的增加,像素间的相关性有所增加,估计有较高的压缩率。压缩率(CR)定义为

$$CR = \frac{\text{原图像比特数}}{\text{压缩图像比特数}} \quad (1-1)$$

这种像素间相关性的增加意味着,例如如果取样速率增加 2 倍,那么压缩图像所需的比特数的增加将少于 2 倍(尽管总比特数仍将增加)。

- (2) 当量化电平数增加时,在某种程度上像素间相关性有所减小,因此减少了可实现的压缩。
- (3) 任何源噪声(例如,印相底片的颗粒噪声)或由传感器引入的噪声将减小像素间的相关性,并减小可实现的压缩量。

1.5 图像压缩标准化工作

现今世界唯一广泛利用的数字图像压缩标准是由国际电报电话咨询委员会 (CCITT) 的工作组建立的国际数字传真第 3 组和第 4 组编码标准。这些标准只应用于诸如文本和文件的双电平图像, 不能用于连续色调图像或图像序列。然而, 跨越多种多样的应用领域, 对以数字形式处理这些图像的需求日益增加, 故有必要制定另一些压缩标准。这些标准的采用不仅便于跨越应用领域的图像交换, 而且也有助于很多实时图像压缩系统所需专门硬件的成本大大降低。目前, 图像压缩标准化工作集中于以下三方面:

(1) 双电平图像

在 ISO-IEC/JTC1/SC2/WG8^[1] 和 CCITT SG VIII NIC^[2] 共同主办下, 于 1988 年建立称为 JBIG (为联合双电平图像组) 的委员会, 来从事双电平图像压缩与解压标准的研究。由于 CCITT 第 3 组和第 4 组标准已存在于此范围内, 故该组的最初重点是在其应用范围内 (例如在任一 CCITT 双电平参考图像中), 寻求一种优于现有标准的算法, 并且还要将它的效能推广到其它应用中。例如, 要求标准必须遵循对图像具有逐行表示能力, 这对于具有可变分辨率和/或质量要求的很多显示或传输设备来说, 在有可能进行数据压缩的环境中, 它是必不可少的性能。此外, 该标准算法应是自适应的, 因此它能有效地压缩具有不同性能的图像。例如, 当利用现有的传真标准压缩数字化半色调图像时, 由于当前标准利用基于二元文件和文本统计的预先确定的编码表, 所以可造成 50% 的扩充数据。JBIG 提出的算法成功地压缩了这种数据。

[1] 国际标准化组织, 连接技术委员会 1, 字符集和编码分会, 第 8 研究组 (图像与声频信息编码表示法)。

[2] 新型图像通信研究组 VIII 专家组。

(2) 连续色调, 静帧, 黑白和彩色图像

为了研究连续色调, 静帧, 黑白和彩色图像压缩与解缩的国际标准, 在 ISO-IEC/JTC1/SC2/WG8 和 CCITT SG VIII NIC 联合主持下, 于 1986 年年底建立了称为 JPEG (为联合摄影专家组) 的委员会。建立 JPEG 主要推动 64kb/s 综合业务数字网络 (ISDN) 的多种业务方式的实现。该委员会的目标是确定各种应用标准, 如光-信息传视系统, 台式出版系统, 图形艺术, 彩色传真, 新闻摄影工作, 医学系统, 以及其它很多领域。虽然上面这些领域没有标准可循, 但 JPEG 成员有坚定信念, 可用共同的、通用图像压缩标准来满足大多数应用的要求。JPEG 提出的标准由三个主要部分组成: 1) 基线系统, 它提供简单与有效的算法, 以满足大多数图像编码应用的要求; 2) 一组扩展系统部件, 诸如逐行建立, 它使基线系统满足广泛应用要求; 3) 用于某种压缩的独立无损方法。

(3) 顺序, 连续色调图像

最近, CCITT 已对 64 至 1920kb/s 比特速率范围的可视电话和可视会议编码算法进行标准化工作。自 1988 年以来, 称为 MPEG (为活动图像专家组) 的标准化组, 也在 ISO-IEC/JTC1/SC2/WG8 赞助下进行研究, 以利用具有 1.0~1.5Mb/s 组合比特速率的数字存贮介质, 来开发活动图像与语音的存贮和恢复的标准。此 MPEG 标准目的是研究一种通用技术, 以便用于电子出版, 导航, 电视信息传视系统, 教育, 游戏, 文娱活动, 可视邮政和电视研讨培训等领域。