

一种新的静止图象压缩编码算法*

黎洪松 全子一

(北京邮电大学无线系 北京 100088)

摘要 本文提出了一种新的静止图象压缩编码算法,即 VQ + DPCM + DCT 算法,并与 JPEG 标准的基本系统进行了比较。实验结果表明,新算法的压缩比有较大提高。

关键词 神经网络,矢量量化,图象编码

1 引言

在图象编码中,量化是一个十分重要的问题,矢量量化(VQ)是一种高效率的压缩编码技术,具有很大的潜力。因此,近年来备受关注和重视,并取得了一些令人鼓舞的新进展^[1]。目前,图象矢量量化的研究主要集中在变换域进行,提出了很多分类算法。值得一提的是文献[2]提出的一种基于人眼视觉特性的离散余弦变换(DCT)矢量量化编码方案。通过对 8×8 DCT变换系数进行视觉分类和矢量量化,得到在压缩比为13.3:1时,峰峰信噪比为34.74dB(针对Lena 512×512 亮度图象),主观评价质量良好。

遗憾的是,矢量量化的研究从理论(构造性定理)到实践至今仍未取得突破性的进展。目前的研究表明,单独使用矢量量化方法进行图象压缩(空间域或变换域),其效果并不明显,在高压缩比下,会产生人眼易察觉到的失真(如方块效应、细节模糊等),主观评价质量不尽如人意,其编码性能和压缩比还不能与采用自适应DCT变换编码的JPEG(联合图片专家组)标准相比。这是由于实际图象景物太复杂,且图象信源又非常不平稳,仅用数量有限的码矢去匹配千变万化的图象景物,显然是力不从心的。因此研究矢量量化与其它编码技术相结合的图象编码方法是很有意义的。

随着神经网络的再度兴起,人们已认识到神经网络是实现图象矢量量化编码的一种有效方法,各种神经网络芯片的不断推出,为矢量量化技术的广泛应用提供了保证。基于上述原因,本文提出了一种基于自组织神经网络算法的VQ + DPCM(差分脉码调制)+DCT图象编码方案。实验表明,这种新算法非常有效,与JPEG标准的基本系统比较,其压缩比有较大提高。

1993-12-10 收到,1994-06-07 定稿

* 八五攻关项目 HDTV 资助。

黎洪松 男,1963年生,博士,主要兴趣有:神经网络应用、HDTV技术和多媒体通信技术。

全子一 男,1929年,教授、博士生导师,从事HDTV、数字图象处理、数字电视和图象通信等教学和研究工作。

2 JPEG 基本系统

ISO 和 CCITT 于 1986 年底成立了“联合图片专家组”，简称 JPEG，研究静止图象压缩编码算法的国际标准。1987 年 7 月用电视图象进行了不同方案的主观实验测试，从十个候选方案中选出了三个方案，1991 年 1 月进行了最后测试，得到自适应 (8×8) DCT 编码方案(基本系统)最佳的结论。JPEG 经过细致的工作后，于 1991 年 3 月提出了 ISO CD10918 号建议草案^[3]。

JPEG 标准基本系统如图 1 所示，提供顺序扫描重建的图象，实现“信息有损”图象编码，图象主观质量能达到损伤难以觉察。它采用以 DCT 为主的算法： 8×8 像素的 DCT，根据视觉特性设计的自适应量化器，用 Huffman 编码作变字长编码等。

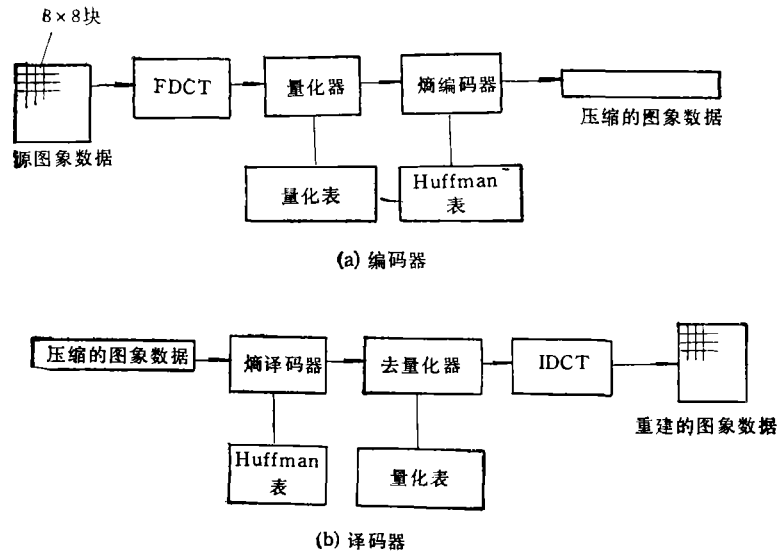


图 1 JPEG 标准基本系统

实验表明，JPEG 基本系统在压缩比为 16:1 左右时，其图象质量很好(彩色 Lena 图象)。

JPEG 基本系统是各国几十年经验之总结，是使静止图象编码技术走向标准化和实用化的一个里程碑。当然，JPEG 的推出并不意味着静止图象编码研究的终止，已提出了一些改进的 JPEG 算法^[4]，目的是提高压缩比和图象质量。随着新兴学科和技术的发展，静止图象编码算法的研究已取得了新的进展^[5]，如 M. Kunt 等人^[6]提出的第二代图象编码技术，分形编码^[7]，小波变换编码^[8]，模型基编码^[9]等。本文则提出了一种基于神经网络的 VQ + DPCM + DCT 编码算法。

3 VQ + DPCM + DCT 算法

JPEG 基本系统是以原始图象的象素值为对象进行 DCT 变换编码来进行数据压缩

的,它并未充分利用图象信源的统计特性和先验知识。在 H.261, MPEG 标准中,由于利用活动图象的时间相关性,通过运动估值可实现帧间预测,从而能实现高压缩比、高质量图象编码。本文基于类似的思想,使用矢量量化方法来预测图象,从而实现高效率的图象压缩,不过未直接涉及时间相关性利用问题。

VQ + DPCM + DCT 编码算法的原理见图 2。其基本方法是,对输入矢量(图象子块) x 先在空间域进行矢量量化,得到量化矢量 \hat{x} ,然后用量化矢量与输入矢量相减,即进行 DPCM,其差值再作 DCT 变换编码,从而实现高压缩比、高质量的图象压缩编码。

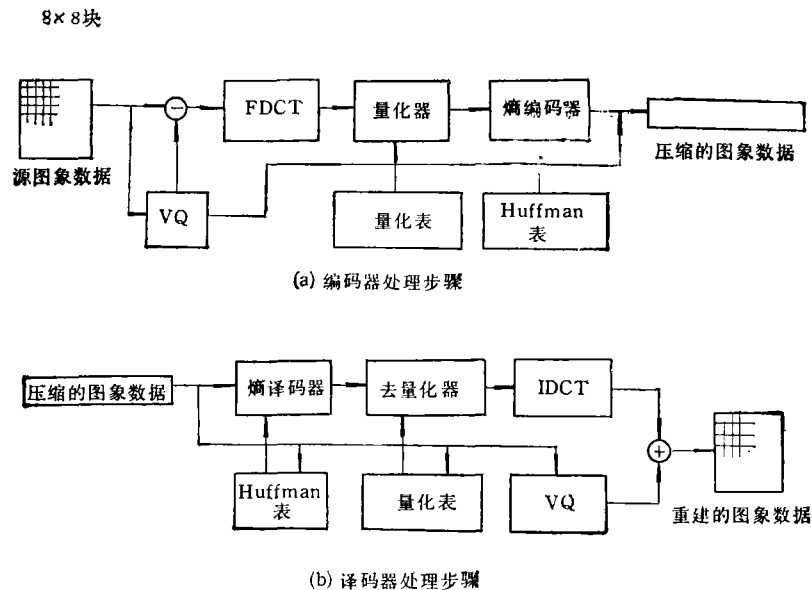


图 2 VQ + DPCM + DCT 算法原理框图

这种新算法能有效地克服矢量量化编码的致命缺点: 图象细节、边缘信息损伤严重和可见的方块效应。方块效应产生的主要原因是各图象子块的交界处像素灰度值失真较大,使边界产生了伪边缘。高频细节模糊则主要由于码书 (Code book) 大小有限,码书的码矢不能很好地匹配细节信息的缘故。不过,若不用 VQ 直接进行图象编码,而是将它作为预测器,则情况就会大不一样。其预测误差通过 DCT 变换编码来补偿,就能有效地减小,甚至消除方块效应和细节模糊。这就是 VQ + DPCM + DCT 编码算法的基本思想。

算法中, VQ 的作用相当于 MPEG 标准中的运动估值器,起预测作用。理论上讲,如果码书是最佳的,它充分反映了图象信源的统计特性,那么量化矢量与输入矢量将非常接近,只需很少的比特来表示其差值信号,其编码系统就能获得较高的压缩比和高质量的编码图象,这正是本文所追求的。

4 码书设计

矢量量化的核心问题是码书设计。目前广泛采用的 LBG 算法^[10],存在以下缺点: 一

是对初始码书很敏感,初始码书的设计是 LBG 算法的关键;二是该算法是一种批处理算法,每次迭代需要处理所有的训练数据,缺乏灵活性和自适应性;三是训练时间长,存在无效码矢。近年来,神经网络开始用于矢量量化器的码书设计^[11,12]。文献[13]提出了一种频率敏感自组织特征映射算法,并得到了频率敏感自组织特征映射算法优于 Kohonen 自组织特征映射算法的结论。码书训练过程如下:

(1) 初始化码书 $W_i(0)$, 设定初始邻域 $NE_i(0)$ 大小和各码矢计数器 $C_i = 1, i = 1, 2, \dots, N, N$ 为码书大小。

(2) 输入训练矢量 $X(t)$ 。

(3) 计算输入矢量和码书中各码矢的失真,失真测度采用均方误差,即

$$d_i(t) = f(C_i) \|X(t) - W_i(t)\|^2, \quad (1)$$

式中 $f(C_i)$ 为频率敏感函数,随 C_i 的增加而单调增大。

(4) 选择最小失真的码矢 i^* ,按下式调整码矢 i^* 和 i^* 的拓扑邻域 $NE_{i^*}(t)$ 内的码矢,即

$$W_i(t+1) = \begin{cases} W_i(t) + \alpha(C_i)[X(t) - W_i(t)], & i \in i^*, NE_{i^*}(t); \\ W_i(t), & \text{其它}; \end{cases} \quad (2)$$

式中 $\alpha(C_i)$ 是学习速率,随 C_i 的增加而单调下降,且 $0 < \alpha(C_i) < 1$ 。 $\alpha(C_i)$ 是一个非常重要的参数,它决定一次权调整的变化幅度,即收敛的快慢程度。

(5) 对所有的训练矢量,重复(2)至(4)。

5 系统分析

5.1 差值信号分析

输入矢量与预测矢量的差值矢量 ΔD 为

$$\Delta D = X - \hat{X}. \quad (3)$$

预测矢量 \hat{X} 与输入矢量 X 越接近,则 ΔD 就越小,可能的压缩效果就越明显。图 3 给出了 $512 \times 512 \times 8\text{bit}$ 原始 Lena 图象象素值的概率分布,曲线比较平坦,没有太明显的峰或谷。图 4 是预测误差值的概率分布,曲线具有非常尖锐的峰,差值信号的动态范

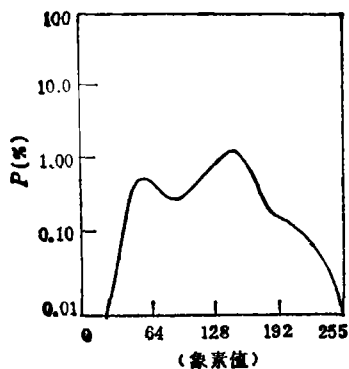


图 3 原始图象象素值概率分布

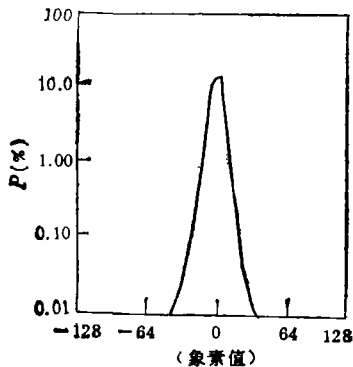


图 4 预测误差图象象素值概率分布

围大大减小。表 1 给出了对应于图 3、图 4 信号的熵值。预测误差的熵值远小于原始图象的熵值,这说明通过预测,减小了信号间的相关性。

5.2 DCT 变换

用 VQ 对输入矢量进行预测后,得到的预测误差仍存在空间结构上的相关性。DCT 变换就是将误差矢量块变换成新的变换系数块。由于大于 8×8 块的变换会陡增矢量量化器的和 DCT 变换器的复杂度,故选择 8×8 块作 DCT 变换。

表 1 对应图 3、图 4 的熵值

项目	参数	熵 值 (bit/pel)
原始图象		7.45
误差图象		4.35

5.3 量化

DCT 变换系数的量化是个非常重要的步骤,表现为:第一,要获得较高的压缩比,通常采用粗量化,而粗量化会使重建图象的质量下降,甚至严重恶化;第二,量化失真是编码系统产生失真的主要原因,且通常是不可恢复的;第三,量化器的设计应考虑人眼的视觉特性。

Lloyd-Max 量化器是一种非线性量化器,由于计算复杂,通常采用均匀量化器。在图象质量要求较高的场合,往往还采用自适应量化方案^[4],即根据图象内容调整量化参数,使量化失真尽可能减小。本文采用基于视觉特性的均匀量化器,量化表由下式产生,

$$q_{ij} = q_0 + (i + j), \quad i, j = 0, 1, \dots, 7, \quad (4)$$

式中 q_{ij} 为量化步长; q_0 为最小量化步长,在自适应量化时可作为调整参数之一。

5.4 熵编码

由原理图 2,熵编码包括对变换系数进行二维熵编码和对矢量量化码矢地址进行 Huffman 编码。

与 JPEG 基本系统不同的是:不再将直流和交流系数分开进行编码,而是直接以零系数游长和非零系数幅度作为事件进行二维 Huffman 编码,码表采用修正的 JPEG 基本系统亮度 Huffman 码表,矢量量化地址码采用 Huffman 编码,其码表略。

6 实验和结果

6.1 实验环境

6.1.1 图象质量评价 图象质量的客观评价采用峰峰信噪比,即

$$\text{PSNR} = 10 \log \frac{255^2}{\text{MSE}} \quad (\text{dB}), \quad (4)$$

式中 MSE 为重建图象的均方误差。

主观评价效果见图 5 中照片。

6.1.2 压缩比计算 设图象的尺寸为 $m \times n$ 个像素,每个像素 B_p bit (亮度),矢量量化地址编码的总比特数为 B_v ,而 DCT 变换编码的总比特数为 B_d ,其压缩比为

$$C_r = \frac{m \times n \times B_p}{B_v + B_d}. \quad (5)$$



(a) 原始 Lena 图象, (b) JPEG 基本系统重建图象, (c) VQ + DPCM + DCT 算法重建图象

图 5 JPEG 基本系统与 VQ + DPCM + DCT 算法比较

6.1.3 码书训练参数 初始码书直接从训练序列中随机选取。

码书大小 $N = 1024$, 训练图象为 $512 \times 512 \times 8\text{bit}$ Lena 亮度图象(图 5(a))。码矢维数 $K = 8 \times 8$ 图象子块。

学习速率由下式确定,

$$\alpha(C_i) = A_1 \exp(-C_i/T_1), \quad (6)$$

式中 A_1 为常数, T_1 为衰减常数。

邻域参数由 $\beta(t)$ 确定,

$$\beta(t) = A_2 + A_3 \exp(-t/T_2), \quad (7)$$

式中 A_2 为最小邻域, 常取 1。 A_3 为训练开始时的最大邻域, T_2 同 T_1 。

频率敏感函数为

$$f(C_i) = C_i^s, \quad (8)$$

式中 s 为频率敏感参数 $0 < s \leq 1$ 。

6.2 实验内容

(1) 码书训练的参数优化。

(2) VQ + DPCM + DCT 算法与 JPEG 基本系统压缩效果比较。

6.3 实验结果

6.3.1 参数优化 图 6 给出了训练参数对 VQ 编码性能(单独 VQ)的影响。

图 6(a) 示出了学习速率 T_1 的影响, 当 T_1 较小时, $\alpha(C_i)$ 的值下降太快, 使码矢的调整幅度变化下降过急, 很难跟得上图象信源的统计特性, 码书性能不尽如人意; 当 T_1 较大时, $\alpha(C_i)$ 值下降很慢, 码矢的调整幅度变化较大, 很容易跳过网络能量函数(均方误差)的全局极小点。最佳学习速率参数, 是使 PSNR 最大的值。

由图 6(b), 当 $A_3(A_2 = 1)$ 较小时, 自组织特征映射网络的效果不明显, 而 A_3 很大时, 其效果不佳, 这是因为在外界刺激最强处(最小失真码矢)所形成的泡(bubble)区的大小是有限的, 离泡区越远, 则神经元间的相互影响就越小, 甚至互相抑制^[5]。

由图 6(c), s 越大, 则频率敏感性越强, 当某个码矢成为响应码矢的次数增加时, 其失真将迅速增大, 从而使该码矢再次成为响应码矢的概率大大降低, 这加大了“公平”的程度。但应该看到, 由于图象信源非常复杂, 有的矢量出现的概率大, 有的则较小, 为了获得高质量的编码图象, 就期望码书尽可能适应图象信源的概率结构, 过份地“公平”会导致码书性能变差。

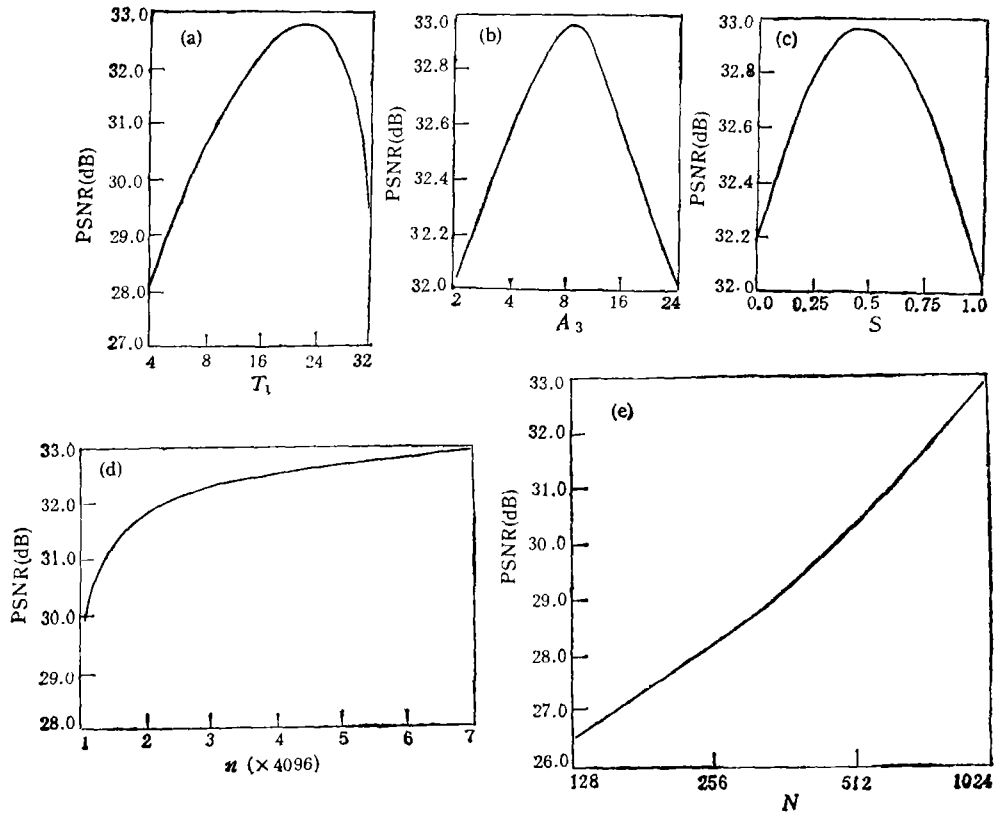


图 6 码书训练参数对 VQ 编码性能的影响

(a) 学习速率对编码性能的影响, (b) 邻域参数对编码性能的影响, (c) 频率敏感参数对编码性能的影响, (d) 训练矢量数目对编码性能的影响, (e) 码书大小对编码性能的影响

由图 6(d), 训练矢量数 n 越大, 其编码性能越好, 但当训练矢量超过一定数目后, 其改善效果并不明显。这说明了两点: 一是训练矢量数目不能太少, 否则码书性能不尽如人意; 二是训练矢量数越大, 则训练时间越长, 因此可适当地限制训练矢量数目, 以缩短训练时间。

由图 6(e), 码书较小时, 编码图象的质量很差, 随码书的不断增大, 其编码图象的质量不断提高。

6.3.2 两种算法比较 表 2 列出了 JPEG 基本系统与 VQ + DPCM + DCT 算法的压缩效果及其比较。图 5(a) 是 Lena 亮度图象的压缩效果。图 5(b) 是 JPEG 基本系统的重建图象, 图 5(c) 是 VQ + DPCM + DCT 算法的重建图象。由表 2 和图 5 看到, VQ + DPCM + DCT 算法明显优于 JPEG 基本系统, 表现在:

第一, 在同样信噪比下, 其压缩比提高了大约 3.3 倍, 这一结果非常令人鼓舞。

第二, 主观效果良好, 即在上述压缩比下, 两种算法的重建图象质量 (主观感觉) 接近, 均无可见失真。

表2 JPEG 基本系统与 VQ + DPCM + DCT 算法压缩效果比较

算法 结果	JPEG 基本系统		VQ + DPCM + DCT 算法		效果改善
	压缩比	PSNR(dB)	压缩比	PSNR(dB)	压缩比
Lena 亮度	9.4	35.78	31.8	35.82	$31.8/9.4=3.38$

7 结 论

本文提出了一种新的静止图象编码算法,它使用矢量量化器作为预测器,充分利用了图象信源的统计特性。这种新算法非常有效,它不仅适用于图象信号,而且也能适用于语音编码。作者将在后续的文章中论述 VQ + DPCM + DCT 算法的硬件实现和用于视频编码。

参 考 文 献

- [1] Nasrabadi N M, King R A. IEEE Trans. on COM, 1978, COM-36(8):957-971.
- [2] Kim D S, Lee S U. IEEE Trans. on COM, 1991, COM-39(4):549-556.
- [3] JPEG digital compression and coding of continuous-tone still images draft, ISO 10918, 1991.
- [4] Wu S W, Gersho A. IEEE Trans. on COM, 1992, COM-40(2):251-254.
- [5] Wallace G K. Commun. ACM, 1991, 34(4):55-65.
- [6] Kunt M, Ikonomopoulos A, Kocher M. IEEE Proc., 1985, 73(4):1102-1115.
- [7] Jacquin A E. IEEE Trans. on IP, 1992, IP-1(1):15-20.
- [8] Devore R A, Jawerth B, Lucier B J. IEEE Trans. on IT, 1992, IT-38(2):37-45.
- [9] Forchheimer R, Eahländer O. Low bit-rate coding through animation, in Proc. Picture Coding Symp. (PCS-87), Stockholm; June 1987.
- [10] Linde Y, Buzo A, Gray R M. IEEE Trans. on COM, 1980, COM-28(1):84-95.
- [11] Fang W C, Sheu B J, Chen O T. A neural network based VLSI vector quantizer for real-time image, in Proc. Data Compression Conf., 1991, 45-49.
- [12] Masrabadi N M, Feng Y. Vector quantization of images based upon the Kohonen Self-Organizing feature maps, in Proc. IJCNN, San Diego, CA: 1988, 101-108.
- [13] 黎洪松. 图像矢量量化-频率敏感自组织特征映射算法: [博士论文]. 北京: 北京邮电大学无线电系, 1994.
- [14] Chen W H, Pratt W K. IEEE Trans. on COM, 1984, COM-32(3):37-43.
- [15] 焦李成. 神经网络系统理论. 西安: 西安电子科技大学出版社, 1991.

A NEW STILL-PICTURE COMPRESSION CODING ALGORITHM

Li Hongsong Quang Ziyi

(Beijing University of Posts and Telecommunications, Beijing 100088)

Abstract A new still-picture compression coding algorithm, VQ + DPCM + DCT algorithm, is proposed. The performance of JPEG baseline system and VQ + DPCM + DCT algorithm is compared. Experimental results show that the new algorithm is very efficient.

Key words Neural network, Vector quantization, Image coding