

基于小波变换和矢量量化的人脸图象压缩

雷寒生 朱小平

(中国科学技术大学计算机系, 中国科学技术大学恒星计算机技术研究有限公司, 合肥 230027)

摘要 在图象的压缩编码中, 矢量量化可以利用某特定类图象(如人脸)的统计特性. 为了在高压缩比下获得较好的压缩效果, 提出了一种新的在小波变换域内进行矢量量化的算法. 该算法用树结构表示小波变换域系数, 并根据各节点值的重要程度, 从每一棵树中提取一个矢量, 然后进行矢量量化; 解码时, 为了使矢量分量能正确地返回到原来树中的正确位置, 需利用 EZW^[1]、SPIHT^[2] 算法的思想标记这棵树, 因为这样才能充分利用父子相关性和兄弟相关性, 从而显著地减少了标记信息. 在提取矢量时, 可用简单的阈值剪枝算法, 也可用 SFQ^[3] 的最佳剪枝算法, 而且后者能进一步提高峰值信噪比. 用该算法对人脸图象进行的压缩试验结果表明, 在高压缩比(100:1左右)下, 恢复的图象质量(视觉效果和峰值信噪比)比通常的小波压缩算法(如 EZW、SPIHT、SFQ 等)好得多. 该算法特别适合于对特定类图象的压缩.

关键词 人脸 小波变换 图象压缩 矢量量化

中图分类号: TP391.41 TN919.81 文献标识码: A 文章编号: 1006-8961(2002)01-0044-06

Facial Image Compression Based on Wavelet Transform and Vector Quantization

LEI Han-sheng, ZHU Xiao-ping

(CS Department, SUN-USTC Institute of Computer Technology Co. Ltd., Hefei 230027)

Abstract In the compression of some particular kinds of image sources, such as human face, vector quantization should naturally be considered to exploit its statistical properties. In this paper, a new vector quantization method in the wavelet transform domain is proposed. We use tree structure to organize coefficients. In each tree, nodes are pruned or retained according to their importance. The survived nodes are serialized to be a vector, which will be the input of vector quantization; a map indicating the positions of these nodes is also stored, which is to be used in decoding. It embeds together the principles of EZW and SPIHT, exploiting fully both the parent-children dependencies and brother dependencies. In our algorithm's framework, SFQ's tree pruning algorithm can also be embedded to increase the PSNR of reconstructed images, though we simply choose the threshold-pruning method to reduce complexity of algorithm. Using the proposed algorithm, the reconstructed facial image in very low bit rate (about 0.08bpp) is superior to that of EZW, SPIHT, SFQ in both perception and PSNR. The algorithm is very suitable for the compression of particular kinds of image sources.

Keywords Human face, Wavelet transform, Image compression, Vector quantization

0 引言

人脸是人类视觉中最常打交道的对象这一, 也是诸如可视电话(Videophone)和会议电视(Telephone)中主要传输的对象. 由于人脸图象数据量巨

大, 因此研究人脸图象的高倍压缩, 是低带宽条件下的可视电话、会议电视进行信息传输的客观要求. 同时, 研究对人脸这一特定类图象的压缩算法, 可以启发人们对其他类图象的压缩.

如今, 基于小波变换的图象压缩研究已经取得了极大的进展, 相比传统的基于 DCT 的图象压缩

(如 JPEG)方法,这种算法不但压缩比例高,而且恢复图象质量也相当好。人们虽从不同角度提出了很多种各有优缺点的小波压缩算法^[4],但由于还没有一种方法算得上是完美的,因此可以说,基于小波变换的图象压缩标准还没有出现。目前,最有希望成为标准的是 SPIHT 算法,该算法从本质上讲,是对 EZW 算法的改进,由于它充分利用了二维图象经过多级小波变换后的特性,算法复杂度低,且具有嵌入性的特点,因此对于通用的图象压缩,不失为是一个很优秀的算法;但是对于特定的图象(如人脸)而言,要取得更好的压缩效果,必须利用图象本身的统计特性。为了充分利用图象本身的统计特性,对小波变换后的系数进行矢量量化,是很自然的想法,因为码本的设计只能更适合某一类图象,而不可能为千变万化的图象设计统一的码本。这样,问题就归结为:如何既能利用小波变换的性质,又能利用某特定类图象的统计特性,来实现图象的高质量压缩。

本文介绍了一种适合于某特定类图象的小波压缩算法,该算法不仅复杂度低,且用它来进行人脸图象的压缩能取得很好的效果,即使在 100 : 1 的高压缩比下,图象仍比较清晰;而用通常的小波压缩算法(如 EZW、SPIHT、SFQ 等)则不能,在 100 : 1 的压缩比下,它们的恢复图象质量已无法容忍。

1 图象小波分解的特点

小波变换就是对二维图象进行整体分解,一次变换后,图象先分为 LL_1, LH_1, HL_1, HH_1 4 个子带,然后 LL_1 又可继续分解为 LL_2, LH_2, HL_2, HH_2 4 部分, LL_2 又可继续分解,依次类推。图 1 是三级变换后的图象划分。观察一幅图象 N 级分解后的图象,可发现一个明显的特点,即相似性, $LL_N, LH_{N-1}, \dots, LH_1$ 相似, $HL_N, HL_{N-1}, \dots, HL_1$ 相似, $HH_N, HH_{N-1}, \dots, HH_1$ 相似,这样,就有了变换域系数的树结构表示: LL_n 中的某一点,对应 LL_{n-1} 中的 4 个点($2 \leq n \leq N$), HL, HH 同理(如图 1 箭头所示)。于是,整个变换域系数(LL_N 除外)由许多树组成(如图 2 所示)。

之所以把系数组合为树结构,是因为在树结构下,由父节点的值就能很好地预测相应的 4 个子节点的值。当父节点值(本文中如不特别申明,值均指绝对值)小于某阈值(threshold) T (如 $T=20$)时,则可能有约 95% 的子节点值也小于 T ^[5],也就是说,父节点

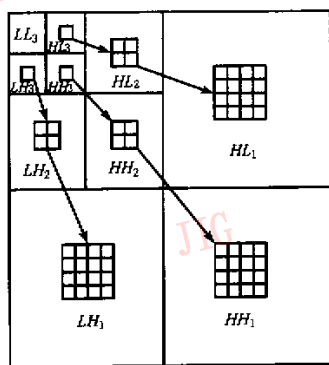


图 1 小波变换示意

(3 级小波分解后图象的划分;箭头表示同一方向系数的父子关系)

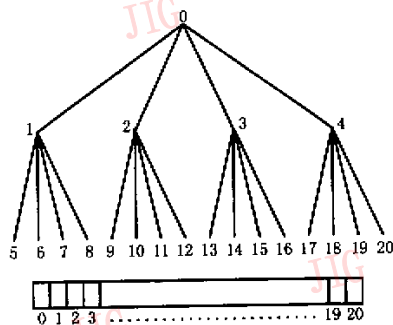


图 2 四叉树示意

(同一方向的具有父子关系的系数组成一棵四叉树;树结构的存储可用顺序数组,存放顺序如标号所示)

值一般大于子节点值,而兄弟节点值又相近。如今,最简单的小波压缩算法是:首先将变换域系数均匀量化,即节点值小于阈值的,就被量化为零,这样就有大量零出现,然后用游程+哈夫曼编码,即能取得很高的压缩比例。很明显,这种算法没有充分利用变换域系数之间的相关性。由树结构体现的父子相关性和兄弟相关性是变换域内的两种主要相关性,另外,还有人提出利用邻居相关性来进行压缩的算法^[6]。但由于同时利用这 3 种相关性是很困难的,因此,本文只利用前两者,实际上,最后一种相关性并不明显。

2 小波压缩的关键问题

图象经小波变换后,变换域内大量的系数值小于阈值 T ,如标准 Lena 图象经 4 级小波变换后,95.2% 的系数值小于 20,即使将这些系数全部设为 0,其恢复的图象质量仍然相当好;如果直接提取出其中的有效系数(值大于或等于阈值 T 的系数),且不

记这些系数的位置信息,则图象的压缩比至少已在20:1以上。但问题是,由于这些系数提取出来后,解码时要能正确地返回原处,因而就必须有个能反映这些有效系数原来位置的map信息。如今,无论哪一种小波压缩算法,因为均必然含有map数据,只不过map信息或者隐藏在算法计算顺序里(如SPIHT),或者单独存在(如EZW、SFQ),所以,小波压缩的关键问题是,如何将变换域的有效系数提取出来,而舍去无效系数(值小于阈值 T 的系数)。这个问题可归结为如下两个方面:(1)怎样提取有效系数,删去无效系数,而且当系数提取出来后,应采取什么样的量化策略;(2)如何尽量减少map信息。

利用树结构来提取有效系数是很好的方法。Xiong运用剪枝算法^[3],首先以率失真为依据,将每一棵树修剪到最佳状态(即在一定的比特率下追求最小的失真),然后对修剪后的树节点进行均匀量化。这里仍然涉及到树结构的表达,即map信息问题。由于Xiong的map信息表达使用了预测方法,因此该算法有一定的复杂度。本文算法中则首先直接用简单的阈值法进行剪枝(也可嵌入SFQ的最佳剪枝算法,以提高峰值信噪比PSNR),且map信息的获取不使用预测方法,而借用EZW算法的思想,并以孤立零点、零树根、正有效系数、负有效系数来标记树中节点(详见本文第3节算法实现);然后,每一棵树的(正负)有效节点值按从上到下、从左到右的顺序串起来作为矢量,由于各棵树经剪枝后剩下的树不一样,这样的矢量长短不一,故本文设定一个统一的长度,不足的添零,多余的切为两个矢量;最后,取恰当的矢量长度,并尽量使一个矢量能容纳一棵树的全部有效系数;接着再进行矢量量化和对量化输出标号进行熵编码。解码时,反过来就行了,根据map信息即可将有效系数恢复到原来树中的正确位置。整个编码算法框架如图3所示。

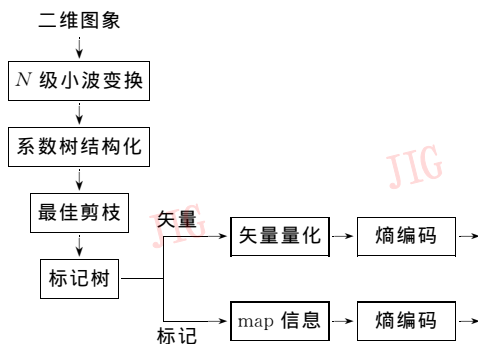


图3 基于小波变换和矢量量化的编码算法流程

李弼程等提出了门限(即阈值)矢量量化方法^[7],其实际上是将所有小于阈值的节点值均设为0,但整棵树不经修剪,只把全部叶子抛弃(即舍弃 LH_1 、 HL_1 、 HH_1),然后量化。这种方法虽简单,且容易实现,但是由于矢量维数大,因此在设计码本时,误差大,计算量也大。本文的算法则是先将有效系数提出,并去除无效系数(包括零树根、零树根的所有子孙和孤立零点),因此维数小,误差小,计算复杂度也低;而且,由于map信息中已包含了正负号,即矢量的分量不再含符号位,所以矢量特征更加集中,码本中可用较少的码矢。

3 算法实现

3.1 N级小波变换

小波变换编码中,滤波器的选择是个有待研究的领域。本文采用双正交小波 $12/4$ 来作为滤波器,实际上,滤波器的选择对编码效果没有明显的影响。另外,变换的级数应根据所期望的压缩比来确定,一般用3~4级,例如,若用4级分解,则 LL_4 是原图象大小的 $1/4^4$,因为 LL_4 中系数最重要,所以,一般将它作有限的无损压缩,再加上其他信息,估计压缩比不会高于 $4^4:1$ 。

3.2 系数数结构化

由于树在未修剪之前都是四叉完全树,因此可直接用一维数组来存储树(见图2),这样既节省空间,又方便操作,还可以随机定位任何节点。本文无论是编码,还是解码,都用一维数组来存储树。

3.3 剪枝算法

剪枝算法可用简单的阈值法来实现:

从下到上,从左到右,检测一棵树的每一个节点,如果节点值小于阈值 T ,则设为零;如果一个节点的所有子孙节点值都为零,则剪去它的所有子孙,该节点即成为叶子节点。实验结果表明,用此阈值法剪枝也能取得很好的效果(算法实例如图4所示)。

如果要追求更高的峰值信噪比,那么剪枝算法可采用Xiong的SFQ算法^[3]。

3.4 标记树

从上到下,从左到右,扫描剪枝后的树,并进行标记:

(1) 如果节点的值负,且绝对值不小于阈值,则输出负有效节点标记NS;

(2) 如果节点的值正,且绝对值不小于阈值,

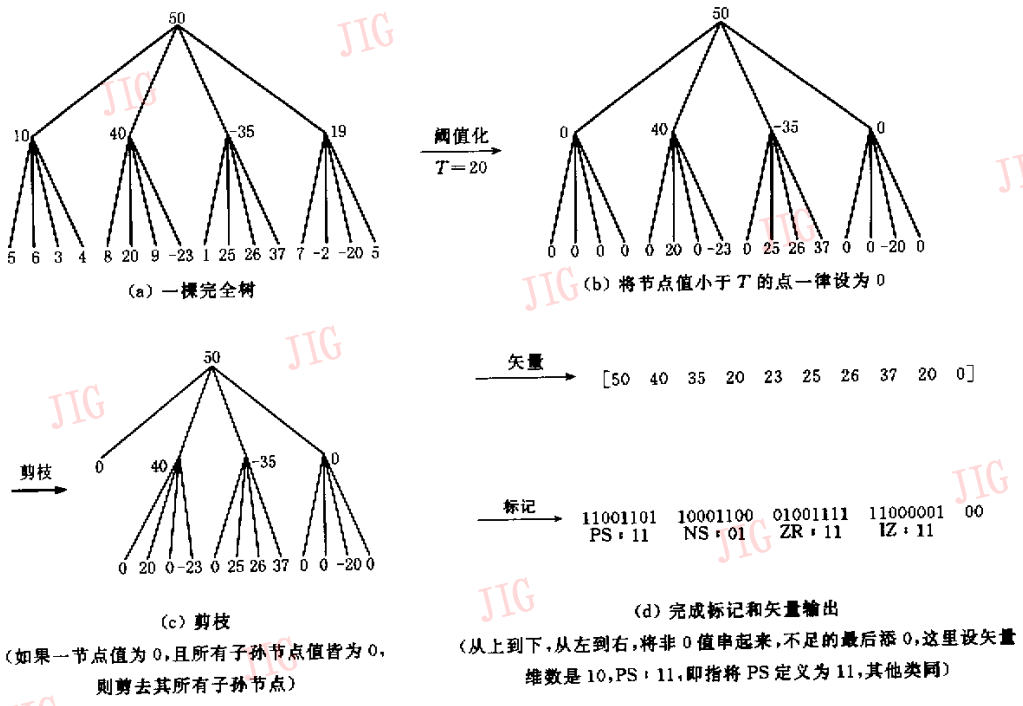


图 4 剪枝和标记算法举例

则输出正有效节点标记 PS;

(3) 如果节点的值为零, 且为叶子节点, 则输出零树根标记 ZR;

(4) 如果节点的值为零, 且不为叶子节点, 则输出孤立零点标记 IZ;

本文把节点标记信息 map 单独存放, 用两个比特正好标记 4 种节点(NS, PS, ZR 和 IZ), 具体的 0、1 指派可视情况而定, 例如, 一般图象的孤立零点都很多, 可设 00 表示孤立零点 IZ; 而将所有有效系数值(绝对值)串联起来当矢量, 再通过量化和编码后单独存放。另外可看到, 如果 3.1 节中采取阈值法, 则编码时, 剪枝算法和标记树可结合在一起完成, 这样算法复杂度将更低。

解码时, 把有效系数取出, 对照节点标记信息 map 可将系数正确恢复到原来树中的位置, 具体操作如下:

(1) 初始化一棵四叉完全树, 设每个节点为“未访问”, 然后从根开始, 从上到下, 从左到右, 设指针 p 指示当前节点。

(2) 取出一个 map 标记:

① 如果是 NS, 则取出一个有效系数, 取反后, 再放入 p 指示的节点, 即成为该节点的值, 然后 p 指示下一个未访问节点;

② 如果是 PS, 则取出一个有效系数, 直接放入

p 指示的节点, 即成为该节点的值, 然后 p 指示下一个未访问节点;

③ 如果是 ZR, 则无须取有效系数, 可直接将零放入 p 指示的节点, 即该节点的值为 0, 再置该节点的所有子孙为“已访问”, 然后 p 指示下一个未访问节点;

④ 如果是 IZ, 则无须取有效系数, 可直接将零放入 p 指示的节点, 即该节点的值为 0, 然后 p 指示下一个未访问节点;

(3) 如果一棵树已全部标记(此时 p 在指示最后节点之后), 则转向步骤(1), 初始化下一棵树, 否则转向步骤(2)。

由于本文总是尽量将一棵树用一个矢量表示, 将不足的添零, 而多余的则切为两个矢量(切为两个矢量的树极少), 因此, 当一棵树未处理完, 但矢量的分量已被取完时, 则应取出下一个矢量; 如一棵树已处理完, 但矢量的分量还没取完, 则抛弃该矢量, 在开始下一棵树标记时, 总是从新的矢量中读取矢量分量作为有效系数。

3.5 码本设计

经典的码本设计算法是 LBG 算法^[8], 而 KSFM、FSCL、CL 也是常使用的码本设计算法。本文则采用神经网络的 CL (competitive learning) 算

法^[9],其在码本设计时是将训练矢量通过竞争性学习,自动归类为 L 个码矢.

本文从 Yale 人脸图象库中随机抽取 15 幅(不包括本文实验用的 Subject1 和 Subject2)人脸图象),经小波变换、剪枝算法提取出矢量后,并把它们作为训练矢量.小波分解的一个重要参数是级数 N ,剪枝算法的一个重要参数是阈值 T ,码本设计的重要参数是矢量维数 D 和码矢个数 L ,实际工作时,可根据图象压缩的不同需要自行设定这 4 个参数.由于本文追求的是高压缩比(约 100 : 1),所以在实验中设 $N=4, T=40, D=10, L=256$.

若以某一类信源为训练对象,则其训练的结果已经融入了该对象的统计特性.至于人脸图象有什么样的具体的统计特性,本文在此不做详细讨论,但本文的算法框架内可结合人脸的特点,在剪枝过程中,可使某些特征(如眼睛,鼻子,嘴巴等)信息尽量少损失或不损失,不过这要依赖于人脸特征部位的正确定位,这是下一步将要研究的问题.

3.6 熵编码

对矢量标号和 map 比特流进行熵编码可进一步压缩数据,而常用的熵编码算法是 Huffman 编码(或自适应 Huffman 编码)、算术编码(或自适应算术编码).一般来说,Huffman 编码速度快,但效果比算术编码稍差;算术编码则相反,因此可根据不同的应用需要来进行选择.本文实验中统一采用算术编码.

4 实验结果

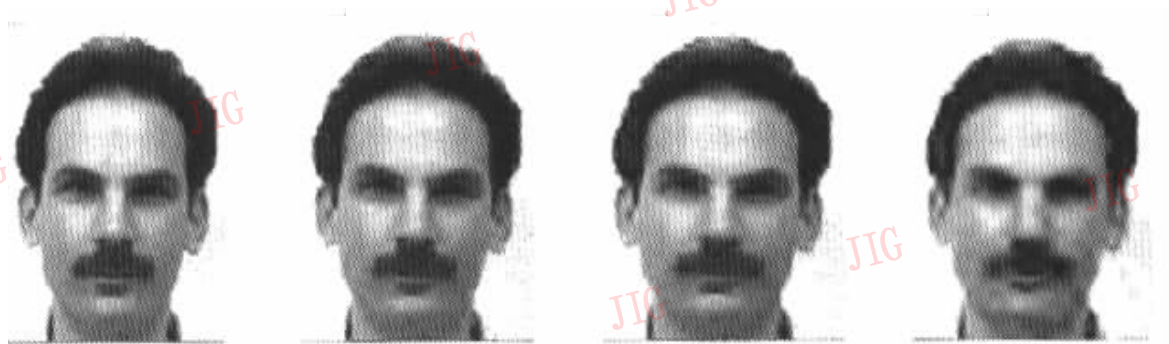
本文用上述方法,对人脸图象进行了压缩实验,

并与 SPIHT 算法进行了比较(SPHIT 算法程序见文献[10]).实验所用训练码矢来自随机抽取的 15 幅 Yale 人脸图象.实验采用的分解级数 $N=4$,阈值 $T=40$,码本长度 $L=256$,维数 $D=10$.图象经 4 级分解后, LL_4 是原图象大小的 $1/256$,这部分系数非常重要,本文对它单独进行熵编码,能压缩 2 倍左右; LH_4, HL_4, HH_4 所有有效系数总和是原图象大小的 $3/256$,这里除零树根外,每个系数节点均对应于一棵树,一棵树对应一个矢量,这样最后输出的矢量标号总和不超过原图象的 $3/256$,且这些标号经过熵编码后,还能压缩 1.5~2 倍,所以除了 map 外,所有有效系数最后均能压缩至 $2/256$ 左右.另外,map 信息经过熵编码能取得很大的压缩比(能达 10 : 1),因为很多树的结构是很相似的.如果增大阈值 T ,则有效系数将减少,且由于更多的节点在零树剪枝时被剪掉,因而 map 信息减少,压缩比提高.

表 1 列出了分别用本文算法和 SPIHT 算法对两幅 Yale 人脸图象压缩的 $PSNR$ (单位: dB)比较.从表 1 中可看出,在中等压缩比例(64 : 1)以下,本文算法效果次于 SPIHT 算法,这主要是因为本文算法在小波分解和矢量训练时,其定位的是高压缩比.实验结果表明,在高压缩比下,本文算法压缩效果无论是从 $PSNR$ 方面,还是从视觉效果方面来看(图 5、6),都明显地比 SPIHT 算法好,这主要是因为本文的目标是对特定的图象进行压缩.从表 1 中亦可看出,本文算法在高压缩比下,压缩恢复图象的 $PSNR$ 值变化很平缓,而 SPIHT 算法压缩恢复图象的 $PSNR$ 则随压缩比的提高而急剧下降,由此可见,本文算法特别适合于对特定类图象的高倍压缩.



图5 Subject1 不同算法压缩结果



(a) 原图象 (b) 本文算法压缩恢复图象 (c) 本文算法压缩恢复图象 (d) SPIHT 算法压缩恢复图象
 (CR=80 : 1, PSNR=27. 88dB) (CR=128 : 1, PSNR=27. 67dB) (CR=80 : 1, PSNR=25. 93dB)

图 6 Subject2 不同算法压缩结果

表 1 本文算法与 SPIHT 算法压缩恢复图象的 PSNR(dB)比较

		32 : 1	64 : 1	80 : 1	100 : 1	128 : 1
Sub1	本文算法	32. 93	30. 24	28. 90	28. 65	28. 42
	SPIHT 算法	33. 82	29. 59	27. 63	26. 16	23. 93
Sub2	本文算法	31. 03	28. 75	27. 88	27. 70	27. 67
	SPIHT 算法	31. 85	27. 18	25. 93	25. 07	23. 31

5 结 论

本文提出的在小波变换域内进行矢量量化的算法,不仅适合于对特定类(如人脸等)图象的压缩,而且在高压缩比(如 100 : 1)下,无论是从 PSNR 还是从视觉效果的角度看,图象重构质量均比 EZW、SPIHT 和 SFQ 算法更优秀。这主要是因为本文算法不仅借鉴了 EZW、SPIHT 算法的思想,还充分利用了变换域系数的父子相关性和兄弟相关性,同时,还利用矢量训练来挖掘特定类图象统计特性的缘故。

对于人脸图象,如果能准确定位人眼等特征部位,那么就可剪枝算法中尽量减少特征部位的信息损失,这样相信图象视觉效果会更好,这是下一步将进行的研究工作。

参 考 文 献

- Shapiro J. Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans. on Signal Processing, 1993, 41(12): 3445~3462.
- Said A, Pealman W A. A new fast and efficient image codec based on set partitioning in hierarchical trees[J]. IEEE Trans. on Circuits Syst. Video Technol., 1996, 6(6):243~250.

- Xiong Z, Ramchandran K, Orchard T. Space-frequency quantization for wavelet image coding[J]. IEEE Trans. on Image Processing, 1997, 6(5):677~693.
- <http://www.icsl.ucla.edu/~ipl/psnr-results.html>.
- 郭田德,高自友.改进的静态图象零树编码算法[J].计算机学报, 1999, 22(7):692~697.
- 李波,汪海.基于小波包变换的分层预测图象压缩算法[J].计算机学报, 1999, 22(7):685~692.
- 李弼程,胡宗云.基于小波变换的图象矢量量化[J].信号处理, 2000, 16(1):32~36.
- Linde Y, Bnzo A, Gray R. An algorithm for vector quantizer design[J]. IEEE Trans. Comm., 1980, 28(1):84~95.
- Ahalt S C, Krishnamurthy A K, Chen P et al. Competitive learning algorithms for vector quantization [J]. Neural Networks, 1990, 3(5):277~290.
- <http://www.cipr.rpi.edu/research/SPIHT/>.

雷寒生 1975 年生,1998 年获青岛海洋大学计算机系学士学位,现为中国科技大学硕士研究生.主要研究方向为图象处理、模式识别、计算机视觉。

朱小平 1954 年生,中国科技大学计算机系高级工程师,科大恒星计算机高技术公司副总经理.主要研究方向为人工智能、知识工程等。