两个优于分裂法的初始码书设计算法

李弼程 文 超 平西建
(解放军信息工程学院信息科学系, 郑州 450002)

摘 要 在矢量量化中, 码书起决定性的作用, 它决定了量化的性能; 一般采用 LBG 算法生成码书, 其中一个关键的技术就是初始码书的选取, 通常认为分裂法效果显著. 该文引入贪婪树生长算法来设计初始码书, 得到了两个优于分裂法的初始码书设计算法, 减少了整个码书训练的运算时间, 提高了码书的性能.

关键词 矢量量化 贪婪树生长算法 分裂法

中图分类号: TN911.73 **文献标识码**: A **文章编号**: 1006-8961(2000)01-0048-04

On the Design of Original Codebooks with Two Algorithms Better Than the Splitting Algorithm

LI Bi-cheng, WEN Chao, PING Xi-jian
(*Department of Information Science, Information and Engineering Institute, Zhengzhou* 450002)

Abstract Codebooks are crucial to vector quantizations (VQs), which determine their performance. LBG algorithm is often used to generate codebooks, where choice of original codebooks is a key technique, and the splitting algorithm is generally considered effective. In this paper, the greedy tree growing algorithm is introduced to design original codebooks, and two better algorithms are derived. Compared to the classical splitting algorithm, the total run time with the proposed algerithm is reduced, and the codebook performance is improved.

Keywords Vector quantization, Greedy tree growing algorithm, Splitting algorithm

0 引 言

从 80 年代开始, 矢量量化[1,2]已经广泛用于图象与声音压缩. 在矢量量化中, 码书起决定性的作用, 它决定了量化的性能. 一般地说, 矢量量化器的码书是采用 LBG 算法[1]来设计的, 先选取初始码书, 通过对训练序列的量化、聚类、迭代, 产生局部最优的码书. 本文讨论基于完全搜索的固定率矢量量化器的码书设计. LBG 算法的具体步骤如下:

(1) 给定初始码书 $Y_N^{(0)}=\{Y_i \mid i=1, \cdots, N\}$, 并置 $n=0$, 设起始平均失真 $D^{(-1)} \to \infty$, 给定计算停止门限 $\epsilon(0<\epsilon<1)$.

(2) 用码书 $Y_N^{(n)}$ 中的码字为形心, 根据最近邻域原则把训练序列 $TS=\{X_r \mid r=1, \cdots, M\}$ 划分为 N 个胞腔: $S_j^{(n)}=\{X \mid d(X, Y_j) \leq d(X, Y_i), i \neq j, Y_i, Y_j \in Y_N^{(n)}\}$, $j=1, \cdots, N$.

(3) 计算平均失真与相对失真. 平均失真为

$$D^{(n)}=\frac{1}{M} \sum_{r=1}^{M} \min_{Y \in Y_N^{(n)}} d(X_r, Y) \tag{1}$$

其中 $X_r \in TS; r=1, \cdots, M$, 距离 $d(\cdot, \cdot)$ 取为欧氏距离.

相对失真为

$$\tilde{D}^{(n)}=\left|\frac{D^{(n)}-D^{(n-1)}}{D^{(n)}}\right| \tag{2}$$

若 $\tilde{D}^{(n)} \leq \epsilon$, 则停止计算, 则当前码书为设计好的最终码书, 否则进行第(4)步.

(4) 计算(2)中得到的各个胞腔的形心, 以这 N 个形心构成码书 $Y_N^{(n+1)}$, 并置 $n=n+1$, 返回到(2).

基金项目: 郑州信息工程学院科研基金项目资助(YP9832063)
收稿日期: 1999-01-19; 收到修改稿日期: 1999-04-19

其中, 胞腔(集合) S 的形心 Y 由下式给出:

$$Y=\frac{1}{|S|}\sum_{X\in S}X \tag{3}$$

这里, $|S|$ 表示 S 中所含元素的个数.

初始码书的选取算法有随机法, 分裂法, 乘积码法等, 其中分裂法效果显著. 下面我们简单地介绍分裂法[1,2], 简记为 SA(Splitting Algorithm), 具体步骤为:

(1) 计算整个训练序列 TS 的形心, 设为 Y_0, 以 Y_0 作为第一个码字;

(2) 选择扰动矢量 δ, 以 $\{Y_0-\delta, Y_0+\delta\}$ 为初始码书, 利用 LBG 算法, 设计仅含有两个码字的码书 $\{Y_1, Y_2\}$;

(3) 以 $\{Y_1-\delta, Y_2-\delta, Y_2+\delta\}$ 为初始码书, 利用 LBG 算法, 设计仅含有 4 个码字的码书, 再加上扰动矢量 δ 以扩大码字的数目. 如此反复, 经过 $\log_2^N$ 次设计, 就得到所要求的有 N 个码字的初始码书.

其中, 扰动矢量 δ 的取法有两种:

(1) $\delta=\alpha\sigma=(\alpha\sigma_1, \alpha\sigma_2, \cdots, \alpha\sigma_K)$, 其中, 比例因子 $\alpha>0$, σ_i 为训练矢量中第 i 个分量的标准方差, $i=1, 2, \cdots, K$, K 为训练矢量的维数.

(2) $\delta=\alpha\theta=(\alpha\theta_1, \alpha\theta_2, \cdots, \alpha\theta_K)$, 其中, 比例因子 $\alpha>0$, θ 为训练矢量的协方差矩阵的最大的特征值所对应的特征向量.

在分裂法中, 对中间码书的每一个码字都要一分为二, 然后利用 LBG 算法形成数目是原码书的二倍的新码书. 这样设计的码书, 其码字均匀或近似均匀地分布在样本空间, 而实际信源是非均匀的, 容易造成有些码字利用率很低. 此外, 这种分裂法的运算工作量较大.

为了克服分裂法的上述缺点, 我们引入树状矢量量化器与可变率矢量量化器设计中的贪婪树生长算法[3,4]来设计初始码书, 得到了两个优于分裂法的的初始码书设计算法, 减少了整个运算量, 避免了空胞腔的产生, 并且提高了码书的性能.

1　基于贪婪树生长算法的初始码书设计

1.1　基本思想

给定训练序列 $TS=\{X_r|r=1, \cdots, M\}$, 矢量的维数为 K, 需要设计具有 $N(\geq 2)$ 个矢量的码书, 其中, $M\gg N$. 首先把训练序列 TS 看成一个全集 S, 以 S 的形心作根节点 v_0. 以 v_0 为父节点, 利用上述的分裂算法, 得到两个矢量 v_1 与 v_2, 称之为子节点; 同时, 把 S 分裂为两个集 S_1 与 S_2, S_i 表示 S 中距离 v_i 最近的矢量的集合, 事实上 v_i 是 S_i 的形心, 其中, $i=1, 2$. 此时, 用矢量 v_1 与 v_2 构成码书去量化所有的训练矢量, 总失真为 $D_1=\sum_{Y\in S_1}d(Y, v_1)+\sum_{Y\in S_2}d(Y, v_2)$.

我们称没有子节点的节点为终端节点. 贪婪树生长算法开始以 v_0 为根节点, v_1 与 v_2 为终端节点(称这次更改为第一次更改), 以后每次只更改一个终端节点, 即把此终端节点改为父节点, 添加其两个子节点为终端节点, 保持其它的终端节点不变, 从而增加一个终端节点, 使得平均失真的减少量与比特率的增量的比值, 即梯度变量

$$\lambda=\frac{\nabla D}{\nabla l} \tag{4}$$

达到最大. 其中, ∇D 表示更改一个终端节点以后平均失真的减少量, ∇l 为增加一个节点后比特率的增量, 单位是比特/矢量.

假设已经进行了 $L-1(L\geq 2)$ 次更改, 得到了 L 个终端节点 $w_1, w_2, \cdots, w_L$, 全集 S 被分裂成 L 个子集合 $T_1, T_2, \cdots, T_L$, w_i 是 T_i 的形心, 用矢量 $w_1, w_2, \cdots, w_L$ 构成码书去量化所有的训练矢量, 总失真 D_{L-1} 近似为

$$D_{L-1}\approx\sum_{i=1}^{L}\sum_{Y\in T_i}d(Y, w_i) \tag{5}$$

现在要分裂 $w_1, w_2, \cdots, w_L$ 中的一个, 记为 $w_i(i=1, 2, \cdots, L)$, 其它 $L-1$ 个节点保持不变. 以 T_i 为训练集, 从 w_i 出发, 利用上述的分裂法把 w_i 分裂为两个子节点 w_{i1} 与 w_{i2}, 并且把 T_i 对应分成两个子集 T_{i1} 与 T_{i2}, w_{i1} 与 w_{i2} 分别是 T_{i1} 与 T_{i2} 的形心. 如果把 w_i 更改为父节点, 添加其两个子节点为终端节点, 保持其它的终端节点不变, 则得到 $L+1$ 个终端节点 $w_1, \cdots, w_{i-1}, w_{i1}, w_{i2}, w_{i+1}, \cdots, w_L$, 全集 S 被分裂成 $L+1$ 个子集合 $T_1, \cdots, T_{i-1}, T_{i1}, T_{i2}, T_{i+1}, \cdots, T_L$, 此时, 用矢量 $w_1, \cdots, w_{i-1}, w_{i1}, w_{i2}, w_{i+1}, \cdots, w_L$ 构成码书矢量化所有的训练矢量, 总失真 D_L 近似为

$$D_L\approx\sum_{1\leq j\leq L, j\neq i}\sum_{Y\in T_j}d(Y, w_j)+\sum_{Y\in T_{i1}}d(Y, w_{i1})+\sum_{Y\in T_{i2}}d(Y, w_{i2}) \tag{6}$$

令

$$R_i=\sum_{Y\in T_i}d(Y, w_i)-\left|\sum_{Y\in T_{i1}}d(Y, w_{i1})+\sum_{Y\in T_{i2}}d(Y, w_{i2})\right| \tag{7}$$

则节点 w_i 分裂前后平均失真的减少量

$$\nabla D = \frac{D_{L-1} - D_L}{M} \approx \frac{R_i}{M} \qquad (8)$$

因为 w_{i1} 与 w_{i2} 分别是 T_{i1} 与 T_{i2} 的形心, 所以当 T_{i1} 与 T_{i2} 非空时

$$R_i \geq 0 \qquad (9)$$

这里我们考虑固定率矢量量化, 比特率的增量

$$\nabla l = \log_2^{L+1} - \log_2^{L} (\text{比特} / \text{矢量}) \qquad (10)$$

此时, 梯度变量

$$\lambda = \frac{\nabla D}{\nabla l} \approx \frac{R_i}{(\log_2^{L+1} - \log_2^{L})M} \qquad (11)$$

上式定义的梯度变量与选取的节点 w_i 有关.

设 $w_1, w_2, \cdots, w_L$ 中使 $\frac{R_i}{(\log_2^{L+1} - \log_2^{L})M}$ 达到最大的节点为 w_{i_0}, 贪婪树生长算法把 w_{i_0} 改为父节点, 添加其两个子节点为终端节点, 保持其它的终端节点不变, 即每增加一个节点要尽可能地减少平均失真. 实际上, w_{i_0} 使 $R_{i_0} = \max_{1 \leq i \leq L} R_i$. 这样更改 $N-1$ 次, 就可以得到 N 个终端节点. 以这 N 个终端节点为初始码书, 利用 LBG 算法进行码书训练.

1.2 算法 1(GTGA-1)

由上面的分析可知, 如果终端节点数小于 N, 则先要分裂所有的终端节点, 再找出分裂前后失真的减少量最大的终端节点, 把该节点改为父节点, 添加其两个子节点为终端节点, 而其它的终端节点保持不变. 当然, 在更改过程中每一个终端节点只需分裂一次, 把分裂的结果存储起来, 以备调用. 我们简称这种先分裂再更改的贪婪树生长算法为 GTGA-1 (Greedy Tree Growing Algorithm), 具体步骤如下:

(1) 计算整个训练序列 TS 的形心, 设为 v_0, 以 v_0 作为根节点;

(2) 采用分裂算法分裂 v_0, 得到子节点 v_1 与 v_2, 以 v_1 与 v_2 为终端节点, 置 $n=2$;

(3) 分裂新添加的两个终端节点, 存储分裂结果;

(4) 找出所有终端节点中分裂前后失真的减少量最大的端节点, 把该节点改为父节点, 添加其两个子节点为终端节点, 而其它的终端节点保持不变, 置 $n=n+1$;

(5) 如果 $n<N$, 则返回到(3); 否则停止.

1.3 算法 2(GTGA-2)

在 GTGA-1 中, 如果终端节点数小于 N 时, 则先要分裂所有的终端节点, 有些终端节点虽然分裂了, 但其分裂结果没有被利用上. 这样, 我们定义(7)式中的第一项为终端节点 i 对总失真的贡献 G_i:

$$G_i = \sum_{Y \in T_i} d(Y, w_j) \qquad (12)$$

在每一次更改以后, 我们可以先求出每一个终端节点对总失真的贡献, 再利用分裂算法对总失真的贡献最大的终端节点进行分裂, 下一次更改就是把该节点改为父节点, 添加其两个子节点为终端节点, 而其它的终端节点保持不变. 这样更改 $N-1$ 次, 就可以得到 N 个终端节点. 以这 N 个终端节点为初始码书, 利用 LBG 算法进行码书训练. 我们简称这种算法为 GTGA-2, 具体步骤如下:

(1) 计算整个训练序列 TS 的形心, 设为 v_0, 以 v_0 作为根节点;

(2) 用分裂算法分裂 v_0, 得到子节点 v_1 与 v_2, 以 v_1 与 v_2 为终端节点, 置 $n=2$;

(3) 计算新添加的两个终端节点对总失真的贡献;

(4) 分裂所有终端节点中对总失真的贡献最大的端节点, 把该节点改为父节点, 添加其两个子节点为终端节点, 而其它的终端节点保持不变, 置 $n=n+1$;

(5) 如果 $n<N$, 则返回到(3); 否则停止.

2 实验结果与结论

我们把 10 幅尺寸为 512×512 的 256 色灰度图象(人的肖像)划分为 4×4 的块, 共得到 163840 个 16 维的矢量, 以这些矢量构成训练序列, 取 $N=256$, 停止门限 $\epsilon = 0.001$. 初始码书采用 SA、GTGA-1、GTGA-2 3 种算法进行设计, 利用 LBG 算法进行码书训练, 采用峰值信噪比($PSNR$)来评价最终码书的性能, $PSNR$ 的定义为:

$$PSNR = 10\lg\left|\frac{255^2}{MSE}\right| \qquad (13)$$

其中, MSE 为量化后每维的平均均方误差失真. 我们用 C 语言编程, 在 Pentium-S CPU 133 上运行. 主要结果列在表 1 中, 其中, 时间包括初始码书设计与 LBG 算法训练的总和, 非训练序列的平均 PSNR 是对不属于训练序列的 10 幅 512×512 的肖像进行统计得出的.

表 1 SA、GTGA-1、GTGA-2 3 种算法的性能比较

初始码书设计算法	时间(s)	训练序列的 $PSNR$(dB)	非训练序列的平均 $PSNR$(dB)
SA	3 090	36.84	36.38
GTGA-1	2 035	37.03	36.65
GTGA-2	2 355	37.03	36.61

此外, 我们把 256× 256 的 256 色灰度图象 Lenna(如图 1 所示) 划分为 4× 4 的块, 共得到 4096 个 16 维的矢量, 以这些矢量构成训练序列, 取 $N=256$, 停止门限 $\epsilon=0.001$. 初始码书采用 SA、GTGA-1、GTGA-2 3 种算法进行设计, 利用 LBG 算法进行码书训练. 分别对 Lenna 进行量化, 其均方误差 MSE、$PSNR$ 如表 2 所示, 视觉效果如图 2—4 所示.

表 2　测试图象 Lenna 的实验结果

初始码书设计算法	SA	GTGA-1	GTGA-2
MSE	64.28	60.35	61.41
$PSNR$(dB)	30.05	30.32	30.25

从表 1、表 2 中得知, 与传统的分裂算法相比, 使用基于贪婪树生长算法的初始码书设计可以使整个码书训练的运行时间减少近三分之一, 并且降低了均方误差, 使码书的性能($PSNR$) 提高了 0.2dB—0.3dB, 此外, 视觉质量也略有改善. 这主要是因为 GTGA-1 与 GTGA-2 是优化算法, 克服了分裂法中码书的码字均匀或近似均匀地分布在样本空间的不足, 使码字的分布与实际信源的非均匀特性相匹配.

总之, 在初始码书设计中, 与分裂法相比, GTGA-1 与 GTGA-2 既具有理论上的完美性又具有实际应用上的有效性.

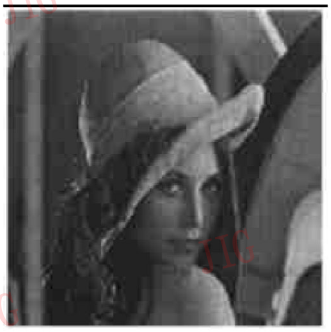

图 1　Lenna(原图)

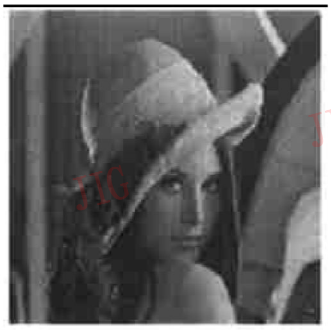

图 2　恢复图象
(SA, $PSNR=30.05$dB)

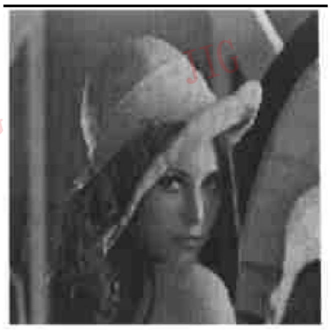

图 3　恢复图象
(GTGA-1, $PSNR=30.32$dB)

图 4　恢复图象
(GTGA-2, $PSNR=30.25$dB)

参 考 文 献

1 Linde Y, Buzo A, Gray R M. An algorithm for vector quantizer design. IEEE Trans Commun, Jan 1980, COM-28(1): 84~ 95.

2 Gersho A, Gray R M. Vector Quantization and Signal Compression. Kluwer Academic Publishers, 1992.

3 John Makhoul, Salim Roucos, Herbert Gish. Vector quantization in speech coding. In: Proc IEEE, 1985, 73: 1551~ 1588.

4 Eve A Riskin, Robert M Gray. A greedy tree growing algorithm for the design of variable rate vector quantizers. IEEE Trans Signal Proc, Nov 1991, 39(11): 2500~ 2507.

李弼程　1998 年毕业于国防科学技术大学, 获博士学位, 现任解放军信息工程学院信息科学系讲师. 主要研究兴趣为小波分析与图象处理. 已公开发表论文 10 余篇.

文　超　现为解放军信息工程学院信息科学系硕士研究生. 主要研究兴趣为图象处理和模式识别.

平西建　1982 年毕业于北京航空学院获硕士学位, 现任解放军信息工程学院信息科学系教授. 主要研究兴趣为图象信源编码理论与方法、图象处理与识别、计算机视觉.