

语音矢量量化编码

章翔凌 李情与

(电子技术系)

摘要 语音矢量量化编码是一项新技术,具有很大的实用价值。本文简要地阐述矢量量化原理之后,着重介绍语音波形矢量量化编码、语音参数矢量量化编码、以及混合形式矢量量化编码,综述近年来国内外学者在这几个方面取得的一些成果,预测它们的发展动向。

关键词 语音信号处理; 矢量量化

1. 引言

数据压缩是提高通信系统效率的一种重要手段。语音信号是通信系统中最普通的信号形式之一,因此语音信号的压缩编码是十分必要的。在正常的讲话速度下,语音的平均信息率为 60bits^{-1} ,而一般的PCM量化体制,信息传递的速率为 64Kbits^{-1} ,故语音信号的压缩是可行的。

用数字信号来近似表示连续信号的过程称之为量化。量化是数据压缩的一种主要的方式,它可分为标量量化和矢量量化两种。由 Shannon 保真度准则下的信源编码定理可知,增加码长可达到速率失真理论的性能限,以及提高通信系统的效率。因而矢量量化的性能较标量量化的性能好。

近年来,矢量量化理论及技术得到了迅速的发展,(见[2,4])。将矢量量化技术应用于语音压缩或编码称为语音矢量量化编码。

本文在简要地阐述矢量量化原理后,着重介绍矢量量化在语音编码中的应用,综述近几年来国内外学者在该方向上取得的一些主要成果,预测未来的发展动向。

2. 矢量量化^[1,2,3,4,5]

矢量量化,简单地说,就是一个具有 k 个样值的有序集(信源矢量)映射为恢复矢量的有限集(码书)中某个矢量(恢复矢量、码字)的映射,而仅用一个二进制数字来表征码书中哪个码字表示原输入矢量。矢量量化编码方式能有效地压缩码率的关键在于,它的整个量化器可视为一编码器(α)和一译码器(β)的组合,信道上所传递的符号是

表征映射关系的二进制数字,而非恢复矢量。

理论上描述矢量量化器可用码书 $A = \{Y_i; i=1, 2, \dots, N\}$ 和空间划分 $S = \{S_i; i=1, 2, \dots, N\}$, 其中 $S_i = \{X \in R^k, q(X) = Y_i\}$, 以及映射 $q: R^k \rightarrow A$, 三个部分合成。记作 $Q(A, S)$ 。

矢量量化器 $Q(A, S)$ 的优劣是以其输入矢量 X 和输出矢量 $Y = q(X)$ 的失真平均值 $D(Q, F) = Ed(X, Y) = Ed(X, q(X))$ 大小而定的, 其中 F 是信源的分布函数。对未知分布的信源而言, D 为时间平均。所谓最优矢量量化器 $Q^*(A, S)$, 就是对于给定矢量维数 k 和码书规模 N , 使得平均失真最小的矢量量化器, 即

$$D(Q^*, F) \leq D(Q, F) \quad (1)$$

码书生成算法是生成最优, 或局部最优矢量量化器的一种算法, 它无疑在矢量量化器研究中起着重要的作用。

LBG算法是码书生成算法中应用较广的一种^[2, 6, 7]。由于语音的分布函数很难用一种解析函数来描述, 故未知分布的LBG算法在语音矢量量化编码的研究中应用较普遍, 论文^[8]详细地分析了这种情况下, 该算法在失真测度不具备凸性等条件下的收敛性问题, 并得到了用该算法生成的矢量量化器极限无偏的条件。

LBG算法实际上是对计算机进行训练的算法, 是聚类算法的一种, 其基本原理是对一已知码书, 找出最佳划分。由优化理论可知, 对不同的初始码书而言, 其收敛的结果亦不同, ^[9]比较了几种初始码书的确定方法, 得出了用其中新提出的逐个聚类法确定的初始码书, 能使LBG算法的迭代次数减少, 而且性能亦不差的结论。

理论上分析矢量量化器的性能, 只能在渐近条件下进行。在大的码书规模 N 下, k 维量化器的失真为^[1, 11]:

$$D_r = S(r, k) \cdot N^{-r/k} \left\{ \int_{-\infty}^{+\infty} [p(X)]^{k/k+r} dX \right\}^{k+r/k} \quad (2)$$

$p(X)$ 是 k 维矢量过程 X 的概率密度, $S(r, k)$ 决定于在 r 阶范数失真测度条件下, k 维空间的划分。它独立于概率密度 $p(X)$ 。

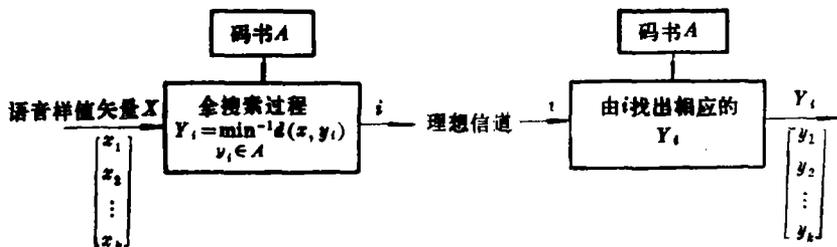
矢量量化器的编码器的编码过程中, 遇到了如何最快, 最准确地搜索到与输入矢量最匹配的恢复矢量的问题。全搜索矢量量化器能最准确地查找到所要求的恢复矢量。如果采用欧氏距离作为失真度量, 一个 k 维 N 级全搜索矢量量化器的编码计算复杂度为: 乘法 Nk 次, 加法 $Nk + N(k-1)$ 次, 比较 $(N-1)$ 次。这样的计算复杂度, 在 N 大于 1024, k 大于 10 的情况下, 现有的技术水平是难以做到实时处理的^[34]。为此人们提出了树搜索矢量量化器, 在性能下降不多以及增加少量存贮复杂度的情况下, 它的计算复杂度较同级数和维数全搜索矢量量化器的要降低许多。另外, 如果构造两个规模为 N_1, N_2 的码书, 让 $N = N_1 \cdot N_2$, 而且两码书中任一对矢量的某种组合能相当于规模为 N 的 k 维码书中的恢复矢量的话, 那么, 在降低一点性能的前提下, 可大大减少矢量量化器的计算和存贮复杂度, 这就是多级矢量量化器和乘积码矢量量化器的出发点。

3. 语音矢量量化编码

根据输入矢量的内容不同, 语音矢量量化编码可分为波形矢量量化编码, 参数矢量量化编码以及两者的混合。

1) 波形矢量量化编码^[15, 23]

波形矢量量化编码的重要一类是矢量PCM (记为VPCM), 见图1。VPCM是把一组时域连续的语音样值组成一个矢量, 然后对这个随机矢量进行矢量编码。



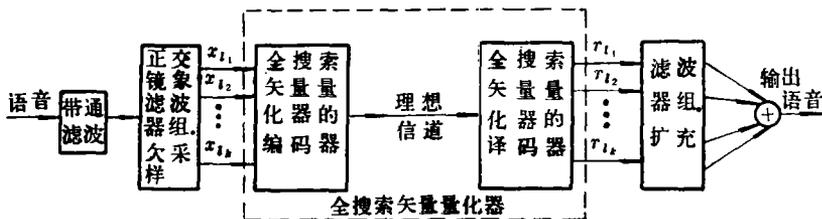
在VPCM中常用的失真度量是均方误差测度。此测度具有计算和分析的方便, 但不能很好地反映主观质量的特性。用它作为失真度量, 得到VPCM的信噪比近似公式是:

$$\text{SNR} = 6.02 \frac{\log_2 N}{k} + C_k \quad (\text{dB}) \quad (3)$$

上式在 $\log_2 N/k$ 较大时是比较合理的, 其中, C_k 是依赖于 k 的常数。

在 $\log_2 N/k$ 一定的情况下, 不同维数的量化器的信噪比差异仅在常数项 C_k 上, C_k 是随着 k 增大而增大的, 故VPCM的信噪比较PCM大。典型的, C_2 比 C_1 大3dB, 而 C_8 要比 C_1 大7dB。然而, 信噪比的提高是以增大量化器的复杂度为前提。

如果称上面一类的编码方案为时域波形矢量量化编码方案, 那么下面介绍的子带矢量量化器就可视为频域波形矢量量化编码方案了。子带编码的量化噪声包含在频带内, 可以防止外频率范围内的量化噪声对该频率范围的掩蔽。在每一频带内可以采用单独的自适应量化阶距, 可调节量化噪声随频率的分布形式, 它能尽量保留在低频段的基音及特征频率的结构, 而在较高频段, 又能较好地恢复摩擦和清音。子带矢量量化器可以充分发挥子带编码的这些长处。子带矢量量化器如图2所示。在图2中, k 维矢量由语谱中每一个子带取一个样值来形成, 故矢量中的元素具有某些能反映频段属性的特征。通



过矢量量化,就能很好地利用语谱中不同频段样值的相关性。

2) 参数矢量量化编码^[17,18,19]

参数矢量是指LPC参数,故参数矢量量化编码亦称为LPC矢量量化编码。它是矢量量化用于语音压缩系统最成功的例子之一。LPC矢量量化编码是把一帧语音通过线性预测分析得到的一组预测系数(在用Itakura-Saito失真度量时,亦包括误差能量),作为一个矢量来进行量化编码的。

LPC分析见图3。它可分为两个部分:第一部分是辨识;第二部分是压缩或量化。在第一部分中,找出一个最匹配输入语音帧的全极点模型。这里常用Itakura-Saito或似然比失真度量^[1,2,10,17,19]。而在第二部分中,却是对每个预测系数和误差能量分别独立进行标量量化,量化所用到的失真测度往往和前面的二种不吻合。

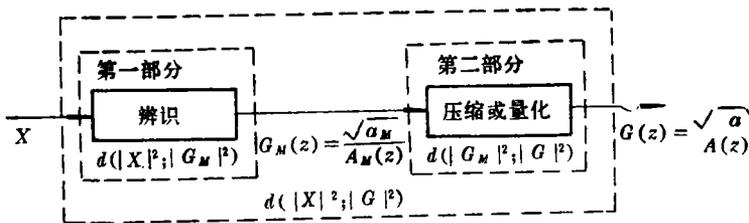


图3 LPC分析的原理框图

在LPC矢量量化编码中,第二部分却是根据Itakura-Saito失真度量对预测系数和误差能量,或用似然比失真度量对预测系数组成的矢量进行量化编码。由于Itakura-Saito失真度量具有三角等式性^[17],故可以把分析失真和量化失真相加。第一部分辨识和第二部分压缩的最优,就等于系统的最优。似然比失真度量虽不具有该性质,但它的计算和误差能量无关,可把预测系数单独组成矢量来量化,致使量化的码书规模降低。同时,它也有分析和量化两过程都使其最小的一致性,因而似然比失真度量在LPC矢量量化编码中应用较多^[18,19,20]。

3) 混合形式矢量量化编码^{[2][24]}

语音混合编码方式,既能克服参数编码中基音提取的困难,获得较好的合成语音质量,又能以低于波形编码速率进行传输。

混合形式矢量量化器的编码器的框图如图4所示,它可视为一个多级矢量量化器。

一帧语音进行LPC分析,得到一组语音预测系数,对这组语音预测系数进行矢量量化编码,形成逆滤波器 $\hat{A}_i(z)$,然后,将这帧语音通过 $\hat{A}_i(z)$ 得到误差信号,再将这误差信号进行矢量量化编码。两次矢量量化得到的信道传递符号一起送到译码器进行译码。由于该矢量量化器是一个多级矢量量化器的形式,故它的计算和存贮复杂度较低,便于设计和实现。

4. 语音矢量量化编码的发展

语音矢量量化编码的研究很大程度上是随着矢量量化理论的发展而发展的,因而,有必要在此先回顾一下矢量量化理论的发展。

矢量量化并不是一个新近才提出的概念。尽管如此,但在七九年以前,由于受种种

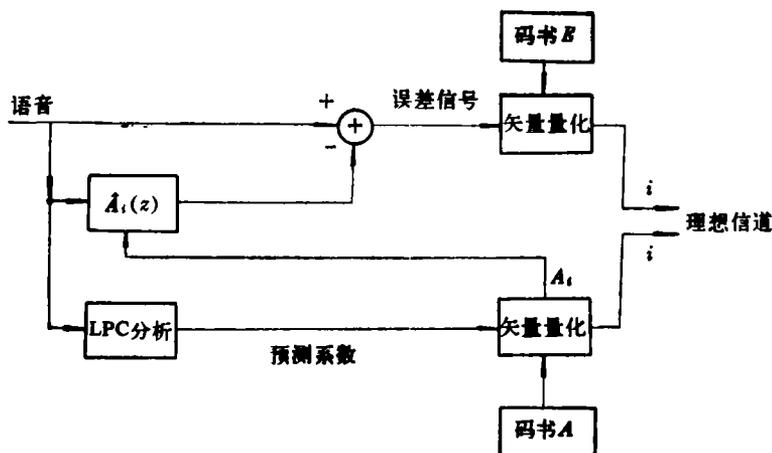


图 4 混合形式矢量量化器的编码器

条件的限制,研究工作始终停留在一个初步的阶段。那时, P. Zador 所完成的多维量化器的渐近性分析, D. J. Newman 所证明的规划六角形阵的最佳性^[5]等等,是比较引人注目瞩目的。自七九年以后,由于许多新的矢量量化器的提出解决了如何生成最优矢量量化器等一系列问题,使得矢量量化研究有了长足的发展。在渐近性分析方面, A. Gersho 发展了原有的矢量量化渐近分析的理论方法^[11], Y. Yamada 等人又分析了在不同失真测度情况下矢量量化器所具有的渐近性。有关矢量量化器的一些理论性文章汇集于文献^[5]。这一时期最有影响的工作是: Y. Linde, A. Buzo 和 R. M. Gray 三人提出矢量量化器码书生成算法 (LBG 算法)^[7]。它是由 S. P. Lloyd 提出的标量最佳 PCM 设计算法^[5]推广而来的,该算法能得出一个在最小平均失真意义下的局部最优矢量量化器。文^[6]对最优矢量量化器性质等进行了分析,在信源渐近均值平稳和训练序列足够长的条件下,用未知信源分布 LBG 算法生成的矢量量化器和已知分布 LBG 算法生成的具有一致性^[1]。文^[8]对更广泛的失真测度,未知分布 LBG 算法的收敛性进行了研究。另外,人们还提出了一些其它码书设计算法^[27,30,31],不过它们大都是由 LBG 算法派生出来的。近年来,在降低量化器复杂度方面,做了大量的工作。这方面又分成构造新的量化器结构和发现新的搜索算法。关于前者,提出了树搜索矢量量化器^[3,13],格型矢量量化器^[5],多级矢量量化器^[25],乘积码矢量量化器^[27],矩阵量化器^[28],有限状态矢量量化器^[30,31]以及降维矢量量化器^[32]等等。另外,还大量地研究了所谓嵌入矢量量化器结构形式,考虑这个问题的初衷是把矢量量化技术和其它压缩技术相结合,以获得较好的性能。具有这一类结构的矢量量化器有:子带矢量量化器^[23],自适应差分矢量量化器^[26],预测矢量量化器^[28]等等。关于后者,总的来说进行得比较少,主要有几种搜索方法,如投影法、超立方体法、部分距离法^[14]和粗细判决法^[15,16]。

语音矢量量化编码的发展是从参数矢量量化编码开始的。A. Buzo 等人首先提出了参数矢量量化编码的概念^[17]。在此之前, R. M. Gray 等人详细地研究了几种谱失真度量^[10],为以后的工作奠定了基础。文^[17]运用 Itakura-Saito 失真度量,使得整个参数矢量量化器在该度量下达到系统最优。由于这种情况下,码书设计受输入语音幅度的影

响,欲使码书具有广泛的适应性,必将增大其规模,为此,该文提出了次最佳的乘积码码书设计的想法。继后,D.Y.Wong等人为了解决码书和语音幅度无关的问题,采用了似然比失真度量,并采用了树搜索矢量量化^[18,19],以此来降低量化器的编码计算复杂度。他们还在后续的论文中,发表了对基音和误差能量的编码方式的研究成果。这样整个参数矢量量化编码的研究就比较完善。此后,S.Rovcos等人利用语音帧间的相关性,得到极低速率的分段矢量量化声码器^[20]。D.B.Paul采用较复杂的感知失真测度,对自适应矢量量化声码器进行了研究^[21],这方面当然还做了许多工作。采用帧重复技术、并对基音及增益用格形码技术进行编码的 400bits^{-1} 的LPC矢量量化声码器,不久将要成为商品^[35]。这无疑是对七十年代发展起来的 2400bits^{-1} LPC声码器的一个挑战。

语音波形矢量量化编码的研究,起步较晚^[22],但近几年来发展迅速。在中速率语音编码系统中,由于复杂度及其它方面的原因^[3],语音波形矢量量化编码的研究引起了广泛的注意。如果用简单的VPCM形式,欲获得较好的性能,必须增大量化器的维数。然而这是受限的,一方面,增大维数,在给定传输率的情况下,量化器的复杂度却是指数增长,这就迫使采用其它形式的量化器结构^[1,2,3];另一方面,如果对语音波形矢量直接量化,由于前后两输入矢量在时间上是连续的,而与之相应的两恢复矢量却是独立选取,故难以保证前一恢复矢量的终点和后一恢复矢量的端点在时间上连续,这将产生不连续的输出语音。这种现象在码书规模小时较突出,因为可供选择的矢量较少。语音波形矢量量化器常用的形式有:乘积码矢量量化器^[27],预测矢量量化器^[28],有限状态矢量量化器^[31]等等,如自适应矢量预测编码器,在码率为 16kbits^{-1} 时,它的信噪比能做到 $17-20\text{dB}$ ^[3],这在标量量化情况下是难以想象的。

人们对参数及波形矢量量化编码研究的同时,还对混合矢量量化编码系统作了研究^[2,24]。

语音矢量量化器的硬件实现见^[33,34,35],其中所采用的芯片大都是通用的几种信号处理芯片。

文^[36]就信道干扰对语音矢量量化编码系统的影响作了初步的研究,这对该系统进入实用有一定的意义。

5. 语音矢量量化编码研究的当前动向

纵观语音矢量量化编码研究的整个领域,我们把它目前的动向归纳成五个方面:

1) 矢量量化器的研究

由于矢量量化器的编码计算复杂度和存贮复杂度随维数的增长而指数增长,因此在维数较大时,这两种复杂度将大到不能实现的程度。为了降低其复杂度,探索新的矢量量化器结构和提出新的搜索方法已成为很迫切的工作。目前,这方面的工作受到了人们的重视。另外,码书生成算法和生成具有Robust性的矢量量化器的研究,现有的成果是远不能满足实际的需要。

2) 语音波形矢量量化编码的研究

随着民用数字通信网的发展,中速率传输语音信号势在必行。已有的标量量化技术难以胜任此事。矢量量化技术可否胜任?如何胜任?人们正怀着极大的兴趣研究着这方

面的课题。

3) 提高语音参数矢量量化器质量的研究

应用矢量量化技术于语音参数编码, 所达到的压缩率十分可观, 但最后合成的语音音质却不能尽如人意。如何设计出具有能为人们接受的话音质量, 又有较低的传输率的矢量量化声码系统? 这无疑是一个十分诱人的研究课题。

4) 信道干扰对语音矢量量化编码系统的影响的研究

随着语音矢量量化编码系统的实用化, 人们越来越注意到这个方面工作的意义。

5) 硬件实现方面的研究

这是语音矢量量化编码系统实用化关键的环节。这里存在许多技术问题。目前成果不多。

6. 结 语

语音矢量量化编码的研究是一个相当活跃的领域, 各国都有不少专家学者在从事这项研究。虽然这方面的论文发表数目逐年增加, 但此项研究可谓相当年轻, 还有许多问题需要人们去探索 and 解决。

参 考 文 献

- [1] J. Makhouf, S. Roucos, H. Gish: Vector Quantization in Speech Coding, *IEEE Trans. Acoust., Speech, Signal Processing*, Nov. 1985, pp. 1551—1558.
- [2] R. Gray: Vector Quantization, *IEEE ASSP Magazine*, April, 1982, pp. 4—29.
- [3] A. Gersho, V. Cuperman: Vector Quantization: A Pattern-Matching Technique for Speech Coding, *IEEE Commun. Magazine*, Dec. 1984, pp. 15—21.
- [4] 胡征, 寇卫东: 信源编码的新动向——矢量量化技术, 全国数据压缩会议, 1984年, 长沙。
- [5] *IEEE Trans. Inf. Theory*, Mar. 1982.
- [6] R. M. Gray, J. C. Kieffer, Y. Linde: Locally Optimal Block Quantizer Design, *Information and Control*, 1980, pp. 178—198.
- [7] Y. Linde, ed al: An Algorithm for Vector Quantization Design, *IEEE Trans. Commun.*, Jan. 1980, pp. 84—95.
- [8] 章翔凌, 李情与, LBG 算法收敛性及其性质证明, 全国信号处理会议, 1986年。
- [9] 章翔凌, 李情与、胡东荣: 一种新初始码书设计算法, 87年全国信息论会议。
- [10] R. M. Gray, et al: Distortion Measures for speech Processing, *IEEE Trans. Acoust., Speech, Signal Processing*, Aug. 1980, pp. 367—375.
- [11] A. Gersho: Asymptotically Optimal Block Quantizations. *IEEE Trans.*

- Inf. Theory, July 1979, pp.373—380.
- [12] Y.Yamada, et al; Asymptotic Performance of Block Quantizers With Difference Distortion Measures, IEEE Trans. Inf. Theory, Jan.1980, pp.6—14.
- [13] G.M.Gray, H.Abut, Full Search and Tree Search Vector Quantization of Speech Waveforms, IEEE ICASSP82, pp.593—597.
- [14] D.Y.Cheng, et al, Fast Search Algorithm for Vector Quantization and Pattern Matching. IEEE ICASSP84, pp.9.11.1—9.11.4.
- [15] Y.Tie-Cheng, L.Lu-Zheng, An Approach to Speed up the Encoding Process of Vector Quantization for Speech Waveforms, WASSP86, pp.245-248.
- [16] 章翔凌、李情与: An Approach To Speed up the Search Process in Vector Quantizers, to be published.
- [17] A.Buzo, et al; Speech Coding Based Upon Vector Quantization, IEEE Trans. Acoust., Speech, Signal Processing, Oct.1980, pp.562—574.
- [18] D.Y.Wong, H.Gray Jr., An 800bits/s Vector Quantization LPC Vocoder, IEEE Trans. Acoust., Speech, Signal Processing, Oct.1982, pp.770—780,
- [19] R.H.Juang, et al; Distortion Performance of Vector Quantization for LPC Voice Coding. IEEE Trans. Acoust., Speech., Signal Processing, April 1982, pp294—298.
- [20] S.Rovcos, et al; Vector Quantization for Very-Low-Rate Coding of Speech, Globecom'82, pp.1074—1078.
- [21] D.B.Paul; An 800BPS Adaptive Vector Quantization Vocoder Using a Perceptual Distortion Measure, IEEE ICASSP83, pp.73—76.
- [22] H.Abut, et al; Vector Quantization of Speech and Speech-Like Waveforms, IEEE Trans. Acoust., Speech, Signal Processing, June 1982. pp.423—435.
- [23] H.Abut, S.A.Luse; Vector Quantizers for Subband Code Waveforms, IEEE ICASSP84, pp.10.6.1—10.6.4.
- [24] D.Wolf, et al; A Comparative Study of Various Speech Encoding Schemes Using Vector Quantization, IEEE ICASSP84, pp.10.8.1—10.8.4
- [25] B.H.Juang; Multiple Stage Vector Quantization for Speech Coding, IEEE ICASSP82, pp.597—600.
- [26] V.Cuperman, A.Gersho; Adaptive Differential Vector Coding of Speech, Globecom'82. pp.1092—1096.
- [27] M.J.Sabin, R.M.Gray; Product Code Vector Quantizers for Waveform and Voice Coding, IEEE Trans, Acoust., Speech, Signal Proc-

- essing. June 1984, pp.474—488.
- [28] A.Haoui, D.G.Messerschmitt; Predictive Vector Quantization, IEEE ICASSP84, pp.10.10.1—10.10.4.
- [29] C.Tsao, R.M. Gray; Matrix Quantizer Design for LPC Speech Using the Generalized Lloyd Algorithm, IEEE Trans. Acoust., Speech, Signal Processing, June 1985 pp.537—545.
- [30] M.D.Dunham, R.M.Gray; An Algorithm for the Design of Labeled-Transition Finite-State Vector Quantizers, IEEE Trans. Commun., Jan.1985, pp.83—89.
- [31] J.Foster, et al; Finite-State Vector Quantization for waveform Coding, IEEE Trans. Inf. Theory, May 1985, pp.348—359.
- [32] 章翔凌, 李情与, 胡东荣; Dimensionality-Reducing Vector Quantizers for Speech Waveform and VoiceCoding, ICCT'87.
- [33] B.P.M.Tao, H.Abut, R.M.Gray; Hardwave Realization of Waveform Vector Quantizers, IEEE Journal on Selected Areas in Commun., Mar.1984, pp.343—352.
- [34] G.Davidson, A.Gersho; Application of a VLSI Vector Quantization Processor to Real-Time Speech Coding, IEEE Journal on Selected Areas in Commun., Jan.1986, pp.112—124.
- [35] J.Camody, J.Rothweller; Speech Coding at 800 and 400bits⁻¹. ITT Electrical Commun., June 1985, pp260—265.
- [36] 章翔凌, 李情与; The Effects of Channel Errors in VQ Systems for Speech Waveform and LPC Voice Coding, to be published.

Vector Quantization for Speech coding

Abstract

Vector quantization (VQ) for speech coding is an advanced technique with great practical value.

In this paper a simple description of the VQ's principles is presented. Next, VQ for speech waveform coding, speech parameters coding and their mixed form coding are described respectively. Then, an overview of them during the past few years and the prediction of their developing tendency are presented.

Key words: Speech signal processing; Vector quantization